AFOSR-TR-97-0461

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	19 Dec 96	Final Technical Report 01 Jan 93 - 30 Sep 96

4. TITLE AND SUBTITLE
Optimal and robust control of transition and turbulence in plane channel flow

5. FUNDING NUMBERS
F49620-93-1-0078

6. AUTHOR(S)
Prof. Parviz Moin
Assist Prof. Thomas Bewley

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Stanford University
Department of mechanical Engineering
Stanford, CA 94305

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NA
110 Duncan Avenue, Suite B 115
Bolling AFB, DC 20332-8050

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release, distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Efficient feedback control algorithms based on optimal and robust control theories have been formulated and tested in direct numerical simulations of turbulent channel flow. The optimization technique used is based solely on the equations governing the fluid flow and variations of a mathematical statement of the control objective, with out the heuristic procedures normally used to determine effective flow control algorithms. The control algorithms tested are shown to be extremely effective, and a host of new ideas for the determination of simple, implementable, effective control rules for turbulent flows have been proposed and are currently still under investigation.

The problem of transition control via optimal and robust techniques has also been studied to draw parallels between the linear and nonlinear theories on problems of significant interest in fluid mechanics. Results on this problem have also been quite good and clearly demonstrate how the control theories are related. With this insight, an important extension of the concepts of robust control theory to nonlinear problems has been made.

14. SUBJECT TERMS

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassiified	UL

Optimal and robust control of transition and turbulence in plane channel flow

Final Report
AFOSR Grant No. F49620-93-1-0078

Principle Investigator: Parviz Moin
Research Assistant: Thomas Bewley

Department of Mechanical Engineering, Stanford University

Abstract

Efficient feedback control algorithms based on optimal and robust control theories have been formulated and tested in direct numerical simulations of turbulent channel flow. The optimization technique used is based solely on the equations governing the fluid flow and variations of a mathematical statement of the control objective, without the heuristic procedures normally used to determine effective flow control algorithms. The control algorithms tested are shown to be extremely effective, and a host of new ideas for the determination of simple, implementable, effective control rules for turbulent flows have been proposed and are currently still under investigation.

The problem of transition control via optimal and robust techniques has also been studied to draw parallels between the linear and nonlinear theories on problems of significant interest in fluid mechanics. Results on this problem have also been quite good and clearly demonstrate how the control theories are related. With this insight, an important extension of the concepts of robust control theory to nonlinear problems has been made.

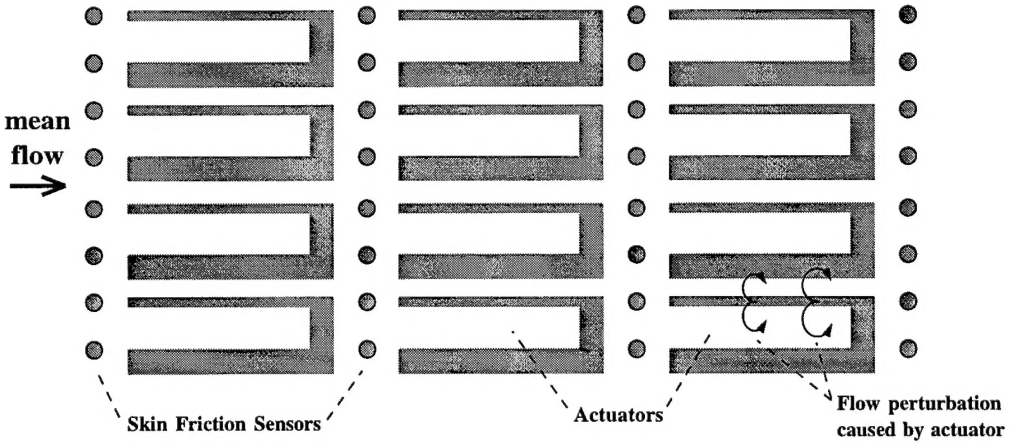
Accomplishments

As a linear "warm-up" project, optimal and robust control techniques were used to effectively control small, two-dimensional, linearly unstable perturbations to a laminar plane channel flow at $Re = 10,000$. The outcome was control rules based on wall-information only which were highly effective at stabilizing the flow system, and is discussed in **Part A: Optimal and robust control of transition**.

The application of control theory to the nonlinear problem of turbulence is, of course, a much greater challenge. The model problem we consider is turbulent flow in a plane channel with blowing and suction distributed over the walls (as an idealization of boundary forcing by discrete MEMS actuators), as illustrated in figure 1. In the case of nonlinear phenomena such as turbulence, iterative approaches must be used based on local linearizations of the flow state. It was also found that optimizations must be performed over finite time intervals which are sufficiently long to accurately reflect the dynamical evolution of the near-wall flow. After some difficulty, an extremely effective implementation of the optimal control technique was tested which reduced the drag of a $Re_\tau = 100$ channel flow by approximately 50%. This far exceeds what is possible using heuristic techniques in the same flow. The calculations also illustrate the sensitivity of important integral flow quantities (such as drag) in a particular flow realization to small modification of the control forcing; thus, a valuable new tool has been developed which may be used to identify coherent turbulent structures responsible for important flow characteristics and, more importantly, where these structures are most sensitive to control forcing. Results are discussed in **Part B: Optimal control of turbulence**.

Comparison of the approaches discussed in Parts A and B led to an understanding of how the concepts from linear robust control theory (*i.e.* $\mathcal{H}_\infty$) may be extended to

(a) Visualization of turbulent channel flow at $Re_\tau = 180$. Shaded regions indicate coherent structures of the near-wall turbulence. Flow is from left to right, walls are dark.



(b) Top view of lower wall, which is covered with sensors and actuators in one possible configuration.

FIGURE 1. The model problem studied in this work is turbulent flow in a plane channel. Small amounts of blowing and suction will be applied through the computational equivalent of closely spaced holes drilled in the walls in response to these turbulent motions in a manner which reduces drag. This is a (very rough) approximation of the eventual physical implementation illustrated in *(b)*.

nonlinear problems in a consistent manner. System robustness is achieved, essentially, by playing the "devils advocate" and attempting to find the "best" feedback control in the presence of a small component of the "worst-case" disturbance forcing the state equation. This type of forcing leads to controls which are less prone to cause instability in the system in the presence of unmodelled disturbances at the price of a slight degradation of performance for the nominal (*i.e.* undisturbed) plant. Such an approach is easily added to the optimal control algorithm discussed in Part B, and is put in a rigorous framework in **Part C: Robust control of turbulence**.

The final, and perhaps most important, result of this project is the development of a technique with which simple, implementable control rules may be rigorously optimized with similar methods. The control rules under consideration in this portion of the work are based on flow information which may be obtained with flush wall-mounted sensors, and determine via simple feedback rules the wall-normal component of the fluid velocity distributed on the walls. The techniques of optimal and robust control theory are used simply to optimize the unknown coefficients in these feedback rules. This work is still under investigation, and is discussed in **Part D: Optimization of practical feedback rules for turbulence control.**

PART A.

Optimal and robust control of transition

Optimal and robust control theories are used to determine feedback control rules that effectively stabilize a linearly unstable flow in a plane channel. Wall transpiration (unsteady blowing/suction) with zero net mass flux is used as the control. Control algorithms are considered that depend both on full flowfield information and on estimates of that flowfield based on wall skin-friction measurements only. The development of these control algorithms accounts for modeling errors and measurement noise in a rigorous fashion; these disturbances are considered in both a structured (Gaussian) and unstructured ("worst case") sense. The performance of these algorithms is analyzed in terms of the eigenmodes of the resulting controlled systems, and the sensitivity of individual eigenmodes to both control and observation is quantified.

---

1. Introduction

The behavior of infinitesimal perturbations to simple laminar flows is an important and well-understood problem. As the Reynolds number is increased, laminar flows often become unstable and transition to turbulence occurs. The effects of the turbulence produced in such flows are very significant and often undesirable, resulting in increased drag and heat transfer at the flow boundaries. Thus, a natural engineering problem is to study methods of flow control such that transition to turbulence can be delayed or eliminated.

Transition often occurs at a Reynolds number well below that required for linear instability of the laminar flow. Orszag & Patera (1983) demonstrate that finite amplitude two-dimensional perturbations can highly destabilize infinitesimal three-dimensional perturbations in the flow. Butler & Farrell (1992) show that the non-orthogonality of the eigenmodes of subcritical flows implies that perturbations of a particular initial structure will experience large amplification of energy before their eventual decay, and suggest that such amplification can sometimes lead to flow perturbations large enough for nonlinear instability to be triggered. Such nonlinear instabilities can lead to transition well below the critical Reynolds number at which linear instability occurs. Results such as these have renewed interest in the control of the small (linear) perturbations, as the mitigation of linear perturbations also lessens the potency of these nonlinear "bypass" mechanisms.

A firm theoretical basis for the control of small perturbations in viscous shear flows is only beginning to emerge. An important step in this direction is provided by Joslin *et al.* (1995) and Joshi, Speyer, & Kim (1996), who analyze this problem in a closed-loop framework, in which the dynamics of the flow system together with the controller are examined. Joslin *et al.* (1995) apply optimal control theory to a problem related to the one presented here; in their approach, the control is determined through an adjoint formulation requiring full flowfield information. Joshi, Speyer, & Kim (1996) consider essentially the same problem analyzed in this paper, and show that a simple constant gain feedback with an integral compensator may be used in a single-input/single-output (SISO) sense to stabilize the flow; a single output (the appropriate Fourier component of

the streamwise drag) is multiplied by some scalar K and summed with a reference signal to determine the corresponding component of the control velocity. This proportional approach is a special case of a class of proportional-integral-derivative (PID) controllers, which combine terms which are proportional, integrals, and derivatives of a scalar output of a system.

The present work extends these analyses to rigorously account for state disturbances and measurement noise. A two-step control approach is used. First, a state estimate is developed from a (potentially inaccurate) model of the flow equations, with corrections to this state estimate provided by (noisy) flow measurements fed back through an output injection matrix L. This state estimate is then multiplied by a feedback matrix K to determine the control. Potentially, this approach can yield better results than a PID controller. In comparison to the PID approach, the present approach has many more parameters in the control law (specifically, the elements of the matrices K and L), which are rigorously optimized for a clearly defined objective. In this manner, multiple-input/multiple-output (MIMO) systems are handled naturally and the controller is coupled with an estimator which models the dynamics of the system itself.

Though a PID approach, such as that of Joshi, Speyer, & Kim (1996), is sufficient to stabilize the present system, it is the authors' judgement that application of modern control theory to the same problem is a timely exercise. Many problems in fluid mechanics, especially those involving turbulence, are dominated by nonlinear behaviour. In such problems, the linear analysis performed in this paper is not valid. However, optimal control approaches, which use full state information, may still be formulated (Abergel & Temam 1990) and performed (Moin & Bewley 1995) with impressive results. In order to make such schemes practical, one must understand how to account for disturbances in a rigorous fashion and how to estimate accurately the necessary components of the state (for instance, the location and strength of the near-wall coherent structures) based on limited flow measurements. The current paper makes these concepts clear in a fluid-mechanical sense, albeit for a linear problem, and thus provides a step in this development.

The controllers and estimators used in this work are determined by application of $\mathcal{H}_2$ and $\mathcal{H}_\infty$ approaches. These techniques have recently been cast in a very compact form by Doyle *et al.* (1989), and are well suited to the current problem, in which the issue of interest is the ability of a closed-loop system to reject disturbances to a laminar flow when only a few noisy measurements of the flow are available. The discussion presented here will involve some tools seen often in the controls literature, such as block diagrams, which are not in common use by the fluids community. Such tools were included only after careful deliberation; it was concluded that these powerful tools are essential in making this development clear, and are thus described in detail when used.

In §2, we derive the governing equations for the present flow stability problem and cast these equations in a standard notation, which makes subsequent application of control theory straightforward. In §3, the control problem is analyzed in terms of the controllability and observability of each individual eigenmode of the system developed in §2. In §4, the control approach developed in Doyle *et al.* (1989) is summarized and applied to the present system. In this control approach, two Riccati equations describe a family of $\mathcal{H}_2$ and $\mathcal{H}_\infty$ controllers, which take into account structured (Gaussian) and unstructured ("worst case") disturbances. Results of these approaches are presented in §5, and §6 presents some concluding remarks.

2. Governing equations

This chapter derives the equations governing the perturbations to a laminar channel flow and casts them in a form to which standard control techniques may be applied. This familiar discussion is presented to precisely define the control problem under consideration. Readers interested only in how the control techniques are applied are advised to proceed directly to §3.

2.1. *Continuous form of flow equations*

Consider a steady plane channel flow with maximum velocity U_0 and channel half-width δ. Non-dimensionalizing all velocities by U_0 and lengths by δ, the mean velocity profile in the streamwise direction (x) may be written $U(y) = 1-y^2$ on the domain $y \in [-1, 1]$. The equations governing small, incompressible, three-dimensional perturbations (u, v, w, p) to the mean flow U are given by linearized Navier-Stokes and continuity

$$\dot{u} + U\frac{\partial}{\partial x}u + U'v = -\frac{\partial p}{\partial x} + \frac{1}{Re}\Delta u \tag{2.1a}$$

$$\dot{v} + U\frac{\partial}{\partial x}v = -\frac{\partial p}{\partial y} + \frac{1}{Re}\Delta v \tag{2.1b}$$

$$\dot{w} + U\frac{\partial}{\partial x}w = -\frac{\partial p}{\partial z} + \frac{1}{Re}\Delta w \tag{2.1c}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \tag{2.2}$$

where $\Delta \equiv \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ is the Laplacian, $Re \equiv U_0\delta/\nu$ is the Reynolds number, ν is the kinematic viscosity, dot ($\dot{}$) denotes $\partial/\partial t$, and prime ($'$) denotes d/dy. A single equation for the normal component of velocity v, found by taking the Laplacian of (2.1b), substituting for Δp from the divergence of (2.1), and applying (2.2), is

$$\Delta\dot{v} = \left\{ -U\frac{\partial}{\partial x}\Delta + U''\frac{\partial}{\partial x} + \Delta(\Delta/Re) \right\} v. \tag{2.3a}$$

The equation for the normal component of vorticity $\omega \equiv \partial u/\partial z - \partial w/\partial x$, found by subtracting $\partial/\partial x$ of (2.1c) from $\partial/\partial z$ of (2.1a), is

$$\dot{\omega} = \left\{ -U'\frac{\partial}{\partial z} \right\} v + \left\{ -U\frac{\partial}{\partial x} + \Delta/Re \right\} \omega. \tag{2.3b}$$

As the domain is homogeneous in the x and z directions, we may Fourier transform the solution such that

$$v(x, y, z, t) = \sum_{k_x, k_z} \hat{v}(k_x, y, k_z, t) \exp[i(k_x\, x + k_z\, z)]$$

$$\omega(x, y, z, t) = \sum_{k_x, k_z} \hat{\omega}(k_x, y, k_z, t) \exp[i(k_x\, x + k_z\, z)].$$

As the various Fourier modes are orthogonal and equations (2.3a) and (2.3b) are linear, the solution for each wavenumber pair (k_x, k_z) is decoupled and obeys the equations

$$\Delta\,\dot{v} = \{-i\,k_x\,U\,\Delta + i\,k_x\,U'' + \Delta(\Delta/Re)\}\, v \tag{2.4a}$$

$$\dot{\omega} = \{-i\,k_z\,U'\}\, v + \{-i\,k_x\,U + \Delta/Re\}\, \omega, \tag{2.4b}$$

where the hat accents ($\hat{}$) have been dropped for notational convenience and the Laplacian now takes the form $\Delta \equiv \partial^2/\partial y^2 - k_x^2 - k_z^2$. Equation (2.4a) is the (fourth order) Orr-Sommerfeld equation for the wall-normal velocity modes, and (2.4b) is the (second order)

equation for the wall-normal vorticity modes. Note the one-way coupling between these two equations. Also note that, from any solution (v, ω), the values of u, w, and p may be extracted by manipulation of the above equations into the form

$$u = \frac{i}{k_x^2 + k_z^2}\left(k_x \frac{\partial v}{\partial y} - k_z\,\omega\right) \tag{2.5a}$$

$$w = \frac{-i}{k_x^2 + k_z^2}\left(k_z \frac{\partial v}{\partial y} - k_x\,\omega\right) \tag{2.5b}$$

$$\Delta p = -2i\,k_x\,U'\,v. \tag{2.5c}$$

Control will be applied at the wall as a boundary condition on the wall-normal component of velocity v. The boundary conditions on u and w are no-slip ($u = w = 0$), which implies that $\omega = 0$ and (by continuity) $\partial v/\partial y = 0$ on the wall.

In this development, it is assumed that an array of sensors, which can measure streamwise and spanwise skin friction, and actuators, which provide wall-normal blowing and suction with zero net mass flux, are mounted on the walls of a laminar channel flow. It is also assumed that a sufficient number of sensors and actuators are installed such that individual Fourier components of wall skin friction and wall transpiration may be approximated, and the analysis is carried through for a particular Fourier mode.

2.2. *Discrete form of flow equations*

The continuous problem described above is discretized on a grid of $N+1$ Chebyshev-Gauss-Lobatto points such that

$$y_l = \cos(\pi l/N) \qquad \text{for } 0 \leqslant l \leqslant N.$$

An $(N+1)\times(N+1)$ matrix $\mathcal{D}$ may be expressed (Canuto *et al.* 1988, eqn. 2.4.31) such that the derivative of ω with respect to y on the discrete set of $N+1$ points is given by

$$\omega' = \mathscr{D}\,\omega \qquad \text{and} \qquad \omega'' = \mathscr{D}\,\omega',$$

where the prime ($'$) now indicates the (partial) derivative of the discrete quantity with respect to y. The homogeneous Neumann boundary condition on v is accomplished by modifying the first derivative matrix such that

$$\tilde{\mathscr{D}}_{lm} = \begin{cases} 0 & l = 0, N \\ \mathscr{D}_{lm} & 1 \leqslant l \leqslant N-1. \end{cases}$$

Differentiation of v with respect to y is then given by

$$v' = \tilde{\mathscr{D}}\,v, \qquad v'' = \mathscr{D}\,v', \qquad v''' = \mathscr{D}\,v'', \qquad \text{and} \qquad v'''' = \mathscr{D}\,v'''.$$

With these derivative matrices, it is straightforward to write (2.4) in matrix form. This is accomplished by first expressing the matrix form of (2.4) on all $N+1$ collocation points such that†

$$\dot{v} = \mathcal{L}\,v \tag{2.6a}$$

$$\dot{\omega} = \mathcal{C}\,v + \mathcal{S}\,\omega, \tag{2.6b}$$

where $\mathcal{L}$, $\mathcal{C}$, and $\mathcal{S}$ are $(N+1)\times(N+1)$. The Dirichlet boundary conditions are explicitly prescribed as separate "forcing" terms. To accomplish this, decompose $\mathcal{L}$, $\mathcal{C}$,

† Note that, for $k_x^2 + k_z^2 \neq 0$, the matrix form of the LHS of (2.4a) is invertible, so the form (2.6a) is easily determined.

and $\mathcal{S}$ according to

$$\mathcal{L} = \begin{pmatrix} * & * & * \\ b_{11} & A_{11} & b_{12} \\ * & * & * \end{pmatrix} \quad \mathcal{C} = \begin{pmatrix} * & * & * \\ b_{21} & A_{21} & b_{22} \\ * & * & * \end{pmatrix} \quad \mathcal{S} = \begin{pmatrix} * & * & * \\ * & A_{22} & * \\ * & * & * \end{pmatrix}$$

where A_{11}, A_{21}, and A_{22} are $(N-1)\times(N-1)$ and b_{11}, b_{12}, b_{21}, and b_{22} are $(N-1)\times 1$. Noting that $\omega_0 = \omega_N = 0$ by the no-slip condition, and defining

$$x \equiv \begin{pmatrix} v_1 \\ \vdots \\ v_{N-1} \\ \omega_1 \\ \vdots \\ \omega_{N-1} \end{pmatrix} \quad A \equiv \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix} \quad B \equiv \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \quad u \equiv \begin{pmatrix} v_0 \\ v_N \end{pmatrix},$$

where x is $2(N-1)\times 1$, A is $2(N-1)\times 2(N-1)$, B is $2(N-1)\times 2$, and u is 2×1, we may express (2.6) in the standard form

$$\boxed{\dot{x} = Ax + Bu.} \tag{2.7}$$

The vector x is referred to as the "state", and the vector u is referred to as the "control".

2.3. *Wall measurements*

We will consider control algorithms using both full flowfield information and wall information only. For the latter case, we will assume that measurements made at the wall provide information proportional to the streamwise and spanwise skin friction

$$y_{m1} = \frac{\partial u}{\partial y}\Big|_{\text{upper wall}} \qquad y_{m2} = \frac{\partial u}{\partial y}\Big|_{\text{lower wall}}$$
$$y_{m3} = \frac{\partial w}{\partial y}\Big|_{\text{upper wall}} \qquad y_{m4} = \frac{\partial w}{\partial y}\Big|_{\text{lower wall}}.$$

Equations (2.5a) and (2.5b) allow us to express these measurements as linear combinations of v and ω. Defining $a \equiv i\,k_x/(k_x^2+k_z^2)$ and $b \equiv -i\,k_z/(k_x^2+k_z^2)$ and using the derivative matrices, the measurements are expressed as

$$y_{m1} = \left(a\,\mathcal{D}\,\tilde{\mathcal{D}}\,v + b\,\mathcal{D}\,\omega\right)_{\text{upper wall}} \qquad y_{m2} = \left(a\,\mathcal{D}\,\tilde{\mathcal{D}}\,v + b\,\mathcal{D}\,\omega\right)_{\text{lower wall}}$$
$$y_{m3} = \left(b\,\mathcal{D}\,\tilde{\mathcal{D}}\,v + a\,\mathcal{D}\,\omega\right)_{\text{upper wall}} \qquad y_{m4} = \left(b\,\mathcal{D}\,\tilde{\mathcal{D}}\,v + a\,\mathcal{D}\,\omega\right)_{\text{lower wall}}.$$

Now decompose $\mathcal{D}\,\tilde{\mathcal{D}}$ and $\mathcal{D}$ according to

$$(\mathcal{D}\,\tilde{\mathcal{D}}) = \begin{pmatrix} d_1 & c_1 & d_3 \\ * & * & * \\ d_2 & c_2 & d_4 \end{pmatrix} \qquad \mathcal{D} = \begin{pmatrix} * & c_3 & * \\ * & * & * \\ * & c_4 & * \end{pmatrix},$$

where c_1, c_2, c_3, and c_4 are $1 \times (N-1)$ and d_1, d_2, d_3, and d_4 are 1×1. Finally, defining

$$y_m \equiv \begin{pmatrix} y_{m1} \\ y_{m2} \\ y_{m3} \\ y_{m4} \end{pmatrix} \qquad C \equiv \begin{pmatrix} a\,c_1 & b\,c_3 \\ a\,c_2 & b\,c_4 \\ b\,c_1 & a\,c_3 \\ b\,c_2 & a\,c_4 \end{pmatrix} \qquad D \equiv \begin{pmatrix} a\,d_1 & b\,d_3 \\ a\,d_2 & b\,d_4 \\ b\,d_1 & a\,d_3 \\ b\,d_2 & a\,d_4 \end{pmatrix},$$

where y_m is 4×1, C is $4 \times 2(N-1)$, and D is 4×2, allows us to express y_m in the standard form of a linear combination of the state x and the control u

$$\boxed{y_m = Cx + Du.} \tag{2.8}$$

The vector y_m is referred to as the "measurement".

2.4. *Inner products and norms*

For this paper, the inner product for two continuous complex functions u and v on the domain $y \in [-1,1]$ is defined by

$$(u,\, v)_\zeta \equiv \int_{-1}^{1} u^*\, v\, \zeta\, dy, \qquad \text{where} \quad \zeta(y) \equiv (1-y^2)^{-1/2}$$

and the star (*) denotes the complex conjugate. The inner product for functions discretized on the collocation points y_l is defined by

$$(u,\, v)_N \equiv \sum_{m=0}^{N} u^*(x_m)\, v(x_m)\, \zeta_m, \qquad \text{where} \quad \zeta_m \equiv \begin{cases} \dfrac{\pi}{2N} & m = 0, N \\ \dfrac{\pi}{N} & 1 \leqslant m \leqslant N-1. \end{cases}$$

For sufficiently smooth functions u, v on a sufficiently large number N of Chebyshev-Gauss-Lobatto grid points (Canuto *et al.* 1988), the discrete inner product approximates the continuous inner product, $(u,\, v)_N \approx (u,\, v)_\zeta$. Note that, for two discrete vectors ξ and η defined only on the interior grid points, or for vectors which are zero at the boundary points $m = 0, N$, the inner product is given simply by

$$(\xi,\, \eta)_N = \frac{\pi}{N}\, \xi^*\, \eta, \tag{2.9}$$

where star (*) applied to a vector denotes conjugate transpose. The norm of v, denoted $||v||$, is defined as the square root of (v, v) for both the discrete and continuous cases. Orthogonality of two functions implies that their inner product is zero.

3. Analysis of control problem

In §2, it was shown that the equations governing small perturbations in a laminar channel flow may be expressed in the standard form

$$\dot{x} = Ax + Bu \tag{3.1a}$$

$$y_m = Cx + Du, \tag{3.1b}$$

where all variables are complex and the system matrix A is dense and non-self-adjoint. We now discuss the eigenmodes of A and identify which of these modes may be modified by the control u and which may be detected by the measurements y_m.

It has been shown (Orszag 1971) that, for $Re \lesssim 5772$, the uncontrolled problem itself is stable and, for $Re > 5772$, weak instability is seen (though most of the eigenvalues

remain stable), with the greatest instability near $k_x = 1.0$ and $k_z = 0.0$. We seek a method to determine the control u which stabilizes the system in a manner which is robust to system uncertainties. To simplify our discussion, we will restrict our attention in the remainder of this work to the particular case $Re = 10,000$, $k_x = 1$, and $k_z = 0$. Joshi, Speyer, & Kim (1996) explore the (Re, k_x, k_z) parameter space further.

For $k_z = 0$ (two-dimensional perturbations), $\mathcal{C} = 0$ in (2.6), entirely decoupling the ω eigenmodes from both the v eigenmodes and from the control $u = (v_0,\ v_N)^T$. In the language of control theory, the ω eigenmodes are thus "uncontrollable" by the control u. (However, it is also seen that the ω eigenmodes are stable, so these modes will, so to speak, "take care of themselves".) Thus, for the remainder of this paper, we will restrict our attention to the the v eigenmodes according to system (3.1) with

$$x = \begin{pmatrix} v_1 \\ \vdots \\ v_{N-1} \end{pmatrix} \qquad A = \begin{pmatrix} & A_{11} & \end{pmatrix} \qquad B = \begin{pmatrix} b_{11} & b_{12} \end{pmatrix} \qquad u = \begin{pmatrix} v_0 \\ v_N \end{pmatrix},$$

where x is $(N-1) \times 1$, A is $(N-1) \times (N-1)$, B is $(N-1) \times 2$, and u is 2×1, and

$$y_m = \begin{pmatrix} y_{m1} \\ y_{m2} \\ y_{m3} \\ y_{m4} \end{pmatrix} \qquad C = \begin{pmatrix} a\,c_1 \\ a\,c_2 \\ b\,c_1 \\ b\,c_2 \end{pmatrix} \qquad D = \begin{pmatrix} a\,d_1 & b\,d_3 \\ a\,d_2 & b\,d_4 \\ b\,d_1 & a\,d_3 \\ b\,d_2 & a\,d_4 \end{pmatrix},$$

where y_m is 4×1, C is $4 \times (N-1)$, and D is 4×2. (All the constituent matrices, vectors, and flow measurements are described in the previous section.)

3.1. *System analysis*

We now address whether or not all of the current system's $N-1$ eigenmodes may be controlled by the $m = 2$ control variables, and whether or not all of these eigenmodes may be observed with the $p = 4$ measurements. To accomplish this, it is standard practice to consider two matrices which characterize the controllability and observability of the system as a whole (Lewis 1995). These are the system controllability Gramian L_c of (A, B) and the system observability Gramian L_o of (C, A), which may be found by solution of

$$A\,L_c + L_c\,A^* + B\,B^* = 0$$
$$A^*\,L_o + L_o\,A + C^*\,C = 0.$$

Note that stable numerical techniques to solve equations of this form, referred to as Lyapunov equations, are well developed (Kwakernaak & Sivan 1972).

If L_c is (nearly) singular, there is at least one eigenmode of the system which is (nearly) unaffected by any choice of control u, and the system is called "uncontrollable". If all uncontrollable eigenmodes are stable, and a controller may be constructed such that the dynamics of the system may be made stable by the application of control, the system is called "stabilizable".

Similarly, if L_o is (nearly) singular, there is at least one eigenmode of the system which is (nearly) indiscernible by the measurements y_m, and the system is called "unobservable". If all unobservable eigenmodes are stable, and an estimator may be constructed such that the dynamics of the error of the estimate may be made stable by appropriate forcing of the estimator equation, the system is called "detectable".

For the present system, the smallest eigenvalue of both L_c and L_o are computed to be near machine zero, indicating that the present system as derived above is both uncontrol-

lable and unobservable. Gramian analysis can not identify *which* of the eigenmodes are uncontrollable or unobservable, however, so it is impossible to predict from this analysis alone whether or not the system is stabilizable and detectable. For this reason, we now develop a method to determine which of the eigenmodes of a system may be affected by the control u and, similarly, which eigenmodes may be discerned by the measurements y_m.

3.2. *Individual eigenmode analysis*

We will now make use of the modal canonical form of the system (3.1) to quantify the sensitivity of each eigenmode of A to both control and observation (Kailath 1980). In order to clarify the derivation, we shall examine each eigenmode of the system separately. Define the eigenvalues λ_i and the right and left eigenvectors, ξ_i and η_i, of A such that

$$\begin{aligned} &\text{right eigenvectors:} && A\,\xi_i = \lambda_i\,\xi_i \\ &\text{left eigenvectors:} && \eta_i^*\,A = \lambda_i\,\eta_i^*, \end{aligned}$$

where the eigenvectors are normalized such that $||\xi_i|| = ||\eta_i|| = 1$ for all i, where $||\cdot||$ is defined in §2.4. Assume A has distinct eigenvalues (this may be verified for the present system described above). Then any x may be decomposed as a linear combination of the (independent but not orthogonal) right eigenvectors such that

$$x = \sum_i \alpha_i\,\xi_i. \tag{3.2a}$$

Differentiating with respect to time,

$$\dot{x} = \sum_i \dot{\alpha}_i\,\xi_i. \tag{3.2b}$$

Also, note that left and right eigenvectors corresponding to different eigenvalues are orthogonal

$$(\eta_j,\,\xi_i) = 0 \qquad j \neq i, \tag{3.3a}$$

but those corresponding to the same eigenvalues are not

$$(\eta_j,\,\xi_j) \neq 0. \tag{3.3b}$$

3.2.1. *Definition of modal control sensitivity*

By (3.1a) and (3.2), we have

$$\sum_i \dot{\alpha}_i\,\xi_i = A\sum_i \alpha_i\,\xi_i + B\,u \tag{3.4}$$

$$= \sum_i \alpha_i\,\lambda_i\,\xi_i + B\,u. \tag{3.5}$$

Taking the inner product with η_j and noting (3.3a) yields

$$(\eta_j,\,\dot{\alpha}_j\,\xi_j) = (\eta_j,\,\alpha_j\,\lambda_j\,\xi_j) + (\eta_j,\,B\,u).$$

By the definition of the inner product (2.9), and noting (3.3b), yields

$$\dot{\alpha}_j = \lambda_j\,\alpha_j + \frac{(B^*\,\eta_j)^*\,u}{\eta_j^*\,\xi_j}.$$

If the vector $B^*\,\eta_j = 0$, then $\dot{\alpha}_j = \lambda_j\,\alpha_j$ for any u. In terms of equation (3.2a), the component of x parallel to ξ_j is not affected by the control u, and the eigenmode is said

to be "uncontrollable". Further, the norm of the coefficient of u

$$f_j = \frac{|\eta_j^* \, B \, B^* \, \eta_j|^{1/2}}{|\eta_j^* \, \xi_j|}, \tag{3.6}$$

which we shall call the control sensitivity of mode j, is a quantitative measure of the sensitivity of the eigenmode j to the control u. Note the dependence of this expression on the matrix $B\,B^*$, which is the same term which drives the Lyapunov equation for controllability Gramian L_c.

3.2.2. *Definition of modal observation sensitivity*

By (3.1b) and (3.2) and assuming, for the moment, that $u = 0$, we have

$$y_m = \sum_j \alpha_j \, C \, \xi_j.$$

If the vector $C\,\xi_j = 0$, then y_m will not be a function of α_j. In terms of equation (3.2a), the component of x parallel to ξ_j does not contribute to the measurements y_m, and the eigenmode is said to be "unobservable". Further, the norm of $C\,\xi_j$

$$g_j = |\xi_j^* \, C^* \, C \, \xi_j|^{1/2}, \tag{3.7}$$

which we shall call the observation sensitivity of mode j, is a quantitative measure of the sensitivity of the measurement y_m to eigenmode j. Note the dependence of this expression on the matrix $C^*\,C$, which is the same term which drives the Lyapunov equation for observability Gramian L_o.

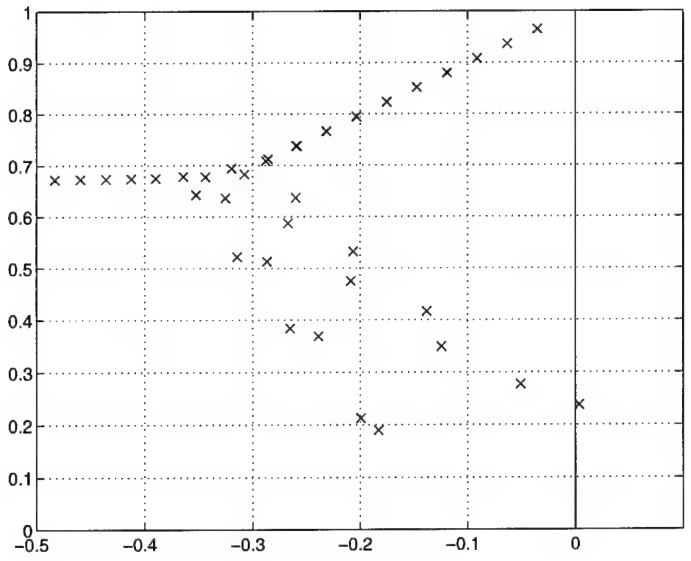
3.3. *Sensitivity of eigenmodes of A to control and observation*

The least stable eigenvalues of A and their corresponding control and observation sensitivities f_j and g_j are tabulated in table 1. Note that the fourth eigenmode is five orders of magnitude less sensitive than the first eigenmode to modifications in the control. In general, those modes in the upper branch of figure 1a (large $|\Im(\lambda)|$) are much less sensitive to control than those in the lower branch (small $|\Im(\lambda)|$). Near the intersection of the two branches ($\Re(\lambda) \approx -0.3$), the control sensitivity is maximum, with this sensitivity decreasing slowly to the left of this intersection ($\Re(\lambda) < -0.3$). It can be predicted that the eigenmodes corresponding to the largest f_j may be affected most upon application of some feedback control u.

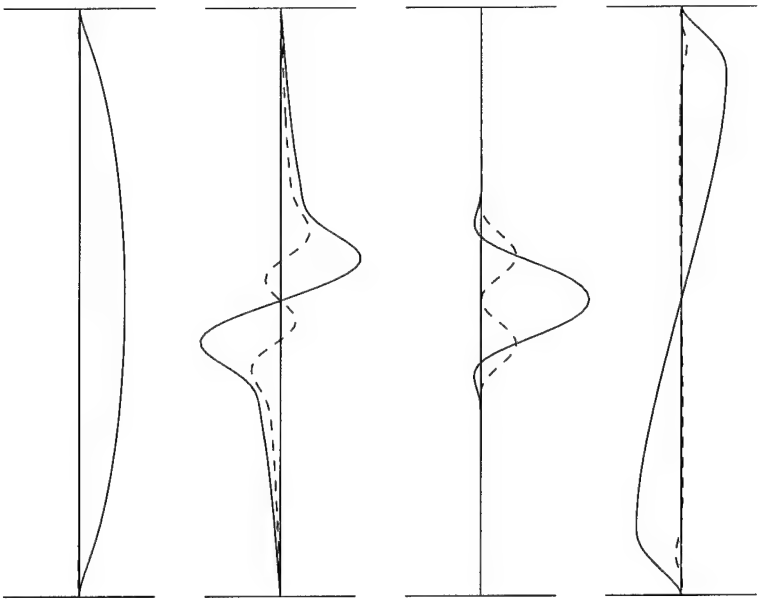
Note that the flow measurements are two orders of magnitude less sensitive to the fourth eigenmode as they are to the first eigenmode. It can be predicted that the state estimates of the eigenmodes corresponding to the largest g_j will be most accurate when estimating the state based on the measurements in the presence of noise.

An important observation from figure 1b is that eigenvalues in the upper branch of figure 1a have corresponding eigenvectors with variations primarily in the center of the channel, and are thus less controllable via wall transpiration and less observable via wall measurements than eigenvalues in the lower branch. This observation is quantified by reduced values of f_j and g_j for these modes in table 1.

The second eigenvalue computed, at $\lambda_2 = -0.0235 + 1.520\,i$ is spurious. Spurious eigenmodes may be easily identified two ways: i) the eigenvalue moves significantly when N is modified slightly, though the eigenvalues reported in table 1 remain converged, and ii) when plotted, spurious modes are dominated by large oscillations from grid point to grid point across the entire domain, though converged eigenmodes are well resolved. Spurious eigenmodes are expected using this approach and may be disregarded.



(a) Least stable eigenvalues: $|\Im(\lambda_j)|$ versus $\Re(\lambda_j)$.



(b) Eigenvectors corresponding to (left to right): $j = 1$ (unstable, lower branch), $j = 3$ (stable, upper branch), $j = 4$ (stable, upper branch), and $j = 5$ (stable, lower branch), plotted as a function of y from the lower wall (bottom) to the upper wall (top). Real component of eigenvector is shown solid and imaginary component dashed. Corresponding eigenvalues are reported in table 1.

FIGURE 1. Least stable eigenmodes of A (no control).

j	λ_j	f_j	g_j
1	$0.00373967 - 0.23752649\,i$	0.266545	102.61
3	$-0.03516728 - 0.96463092\,i$	0.000215	72.85
4	$-0.03518658 - 0.96464251\,i$	0.000005	1.45
5	$-0.05089873 - 0.27720434\,i$	0.026606	347.98
6	$-0.06320150 - 0.93631654\,i$	0.000513	81.39
7	$-0.06325157 - 0.93635178\,i$	0.000021	2.90
8	$-0.09122274 - 0.90798305\,i$	0.000931	83.36
9	$-0.09131286 - 0.90805633\,i$	0.000056	4.32
10	$-0.11923285 - 0.87962729\,i$	0.001587	77.67
11	$-0.11937073 - 0.87975570\,i$	0.000124	5.37
12	$-0.12450198 - 0.34910682\,i$	0.171859	69.50
13	$-0.13822653 - 0.41635102\,i$	0.037660	252.09
14	$-0.14723393 - 0.85124584\,i$	0.002833	63.31
15	$-0.14742560 - 0.85144938\,i$	0.000268	5.59
16	$-0.17522868 - 0.82283504\,i$	0.005581	44.14
⋮	⋮	⋮	⋮
38	$-0.32519719 - 0.63610486\,i$	5.659801	0.78
39	$-0.34373267 - 0.67764346\,i$	4.685315	0.64
⋮	⋮	⋮	⋮
53	$-0.66286552 - 0.67027520\,i$	0.259581	11.58
⋮	⋮	⋮	⋮

TABLE 1. Least stable eigenmodes of A (no control) and the associated control and observation sensitivities. Note that all eigenvalues agree precisely with those reported by Orszag (1971). Calculation used Chebyshev collocation technique with $N = 140$ in quad precision (128 bits per real number). The second eigenmode, which is not shown here, is spurious (see text). Note that the only unstable mode ($j = 1$) for the present system is both sensitive to the control u and easily detected by the measurements y_m.

4. Summary of $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control theories

In §2, it was shown that the equations governing small perturbations in a laminar channel flow may be expressed in the standard form

$$\dot{x} = Ax + Bu \tag{4.1a}$$

$$y_m = Cx + Du, \tag{4.1b}$$

where the constituent matrices A, B, C, and D were summarized and discussed in §3. We now seek a simple method to determine a control u based on the measurements y_m to force the state x towards zero in a manner which rigorously accounts for state disturbances, to be added on the RHS of (4.1a), and measurement noise, to be added on the RHS of (4.1b). Specifically, we will consider feedback of the measurements y_m such that a state estimate $\hat{x}$ is first determined by the system model

$$\dot{\hat{x}} = A\hat{x} + Bu - \hat{u} \tag{4.2a}$$

$$\hat{y}_m = C\hat{x} + Du, \tag{4.2b}$$

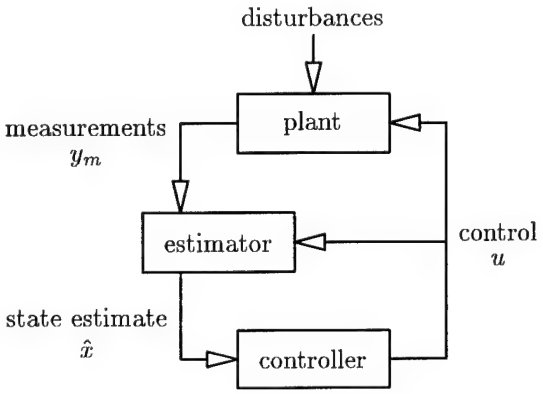
$$\hat{u} = \mathcal{L}(y_m - \hat{y}_m), \tag{4.2c}$$

then this state estimate is used to produce the control

$$u = \mathcal{K}(\hat{x}). \tag{4.3}$$

Equation (4.1), with added disturbance terms on the RHS, is referred to as the "plant", (4.2) is referred to as the "estimator", and (4.3) is referred to as the "controller". The problem at hand is to compute linear time-invariant (LTI) functions $\mathcal{L}$ and $\mathcal{K}$ such that i) the "output injection" term $\hat{u}$ forces the state estimate $\hat{x}$ in the estimator (4.2) towards the state x in the plant (4.1), and ii) the control u computed by the controller (4.3) forces the state x towards zero in the plant (4.1).

The flow of information in this problem is illustrated schematically in the following block diagram.



The plant, which is forced by external disturbances, has an internal state x which cannot be observed. Instead, a few noisy measurements y_m are made, and with these measurements an estimate of the state $\hat{x}$ is determined. This state estimate is then fed back to through the controller to determine the control u to apply back on the plant in order to regulate x to zero.

We will now demonstrate how to apply $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control theories to determine $\mathcal{L}$ and $\mathcal{K}$. (Note that we will redefine several variables used in §2 to derive the Orr-Sommerfeld equation. Considered in the context of this chapter, this should present no confusion.) With this presentation, one set of control equations, involving the solution of two Riccati equations, describes a family of $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control algorithms. The reader is referred to Doyle *et al.* (1989), Dailey *et al.* (1990), and Zhou, Doyle, & Glover (1996) for derivation and further discussion of the control theories summarized here.

4.1. $\mathcal{H}_2$ *control theory*

4.1.1. *Optimal control (LQR)*

The first step in considering the system (4.1) is to consider the problem with no disturbances and measurements which identically determine full information about the state, so that $\hat{x} = x$ (*i.e.* no estimation of the state is necessary). These assumptions are quite an idealization and can rarely be accomplished in practice, but this exercise is an important step to determine the best possible system performance. It is for this reason that the controller in this limit is referred to as optimal. Under these assumptions about the system, the objective of the optimal controller, of the form in (4.3), is to regulate (*i.e.* return to zero) some measure of the flow perturbation x from an arbitrary initial condition as quickly as possible without using excessive amounts of control forcing. Mathematically, a cost function for this problem may thus be expressed as

$$\mathcal{J}_{LQR} \equiv \int_0^\infty \left(||x||^2 + \ell^2 \, u^* u \right) \, dt. \tag{4.4}$$

The term involving $||x||^2$ is a measure of the flow perturbation x integrated over the time period over which this perturbation decays, which is taken as $t \in [0, \infty)$. The term involving u^*u is an expression of the magnitude of the control. These two terms are weighted together with a scalar ℓ^2, which represents the price of the control. This quantity is small if the control is "cheap" (which generally results in larger control magnitudes), and large if applying the control is "expensive". As the state equation is linear, the cost quadratic, and the control objective regulation, this controller is also referred to as a linear quadratic regulator (LQR).

The mathematical statement of the present control problem, then, is the minimization of $\mathcal{J}_{LQR}$. This results in regulation of x without excessive use of control effort. Note that minimization of $\mathcal{J}_{LQR}$ is equivalent to minimization of the integral of z^*z, where

$$z \equiv \begin{pmatrix} Q^{1/2}\, x/\ell \\ u \end{pmatrix}.$$

and where Q is a diagonal matrix with diagonal entries $Q_{jj} = \pi/N$, as required by the definition of the norm in §2.4. In order to arrive at a form which is easily generalized in later sections, define

$$B_2 \equiv B \qquad C_1 \equiv \begin{pmatrix} Q^{1/2}/\ell \\ 0 \end{pmatrix} \qquad D_{12} \equiv \begin{pmatrix} 0 \\ I \end{pmatrix}.$$

For notational convenience, the state equation (4.1a) will be considered as "forced" with a right hand side forcing term r which shall be set to zero, as this regulation problem simply drives the state towards zero without external command input. The state equation (4.1a), the performance measure z, and the state estimate $\hat{x}$ then may be written

$$\dot{x} = A\,x \; + r + B_2\,u \tag{4.5a}$$

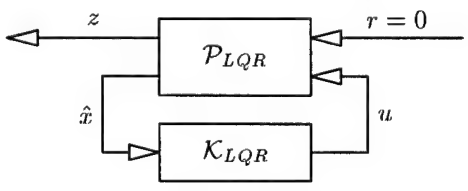
$$z = C_1 x \qquad + D_{12} u \tag{4.5b}$$

$$\hat{x} = \; x. \tag{4.5c}$$

The optimal controller $\mathcal{K}_{LQR}$ is sought to relate the control u to the (precise) state estimate $\hat{x}$. Control is applied to modify the evolution of the state x such that the cost $\mathcal{J}_{LQR}(z)$ is minimized. The important matrices of the system described by (4.5) may be summarized in the shorthand form

$$\mathcal{P}_{LQR} = \begin{array}{c} \\ \dot{x} \\ z \\ \hat{x} \end{array} \begin{array}{c} \begin{array}{ccc} x & r & u \end{array} \\ \left[\begin{array}{c|cc} A & I & B_2 \\ \hline C_1 & 0 & D_{12} \\ I & 0 & 0 \end{array}\right] \end{array}$$

The flow of information is represented by the block diagram



where $\mathcal{P}_{LQR}$ is the flow system given by (4.5) and $\mathcal{K}_{LQR}$ is the optimal controller, which is still to be determined. Note that the command input is $r = 0$ and there are no disturbance inputs; the task of the control u is simply to regulate the state x from nonzero initial conditions back to zero. The state $x = \hat{x}$ is fed back through the controller $\mathcal{K}_{LQR}$ to

control the system. The system output z may be used to monitor the performance of the system.

Given this general setup, a Hamiltonian is defined such that

$$H_2 \equiv \begin{pmatrix} A & -B_2 B_2^* \\ -C_1^* C_1 & -A^* \end{pmatrix}. \tag{4.6a}$$

As shown in Doyle *et al.* (1989), the Hermetian positive-definite solution X_2 to the algebraic Riccati equation defined by this Hamiltonian

$$A^* X_2 + X_2 A - X_2 (B_2 B_2^*) X_2 + (C_1^* C_1) = 0, \tag{4.6b}$$

denoted $X_2 = \mathrm{Ric}(H_2)$, then yields the optimal LTI state feedback matrix

$$K_2 = -B_2^* X_2. \tag{4.6c}$$

The optimal LTI controller $\mathcal{K}_{LQR}$ is then given simply by

$$\boxed{u = K_2 \hat{x}} \tag{4.7}$$

This controller minimizes $\int_0^\infty z^* z \, dt$ in a system with no disturbances and arbitrary initial conditions. Note that standard numerical techniques to solve equations of the form (4.6b) are well developed (Laub 1991).

4.1.2. *Kalman-Bucy filter (KBF)*

When there are disturbances to the system, and thus the state is not precisely known, the state (or some portion thereof) must first be estimated, then the control determined based on this state estimate. The Kalman-Bucy filter, of the form (4.2), accomplishes the required state estimation by assuming that the state disturbances and the measurement noise are uncorrelated white Gaussian processes. To accomplish this, we introduce two zero-mean white Gaussian processes w_1 and w_2 with covariance matrices $E[w_1^* w_1] = I$, $E[w_2^* w_2] = I$, where $E[\cdot]$ denotes the expectation value. With these new disturbance signals, and with G_1 defined as the square root of the covariance of the disturbances to the state equation and G_2 defined as the square root of the covariance of measurement noise, the system (4.1) takes the form

$$\dot{x} = Ax + G_1 w_1 + Bu \tag{4.8a}$$

$$y_m = Cx + G_2 w_2 + Du. \tag{4.8b}$$

The objective of the Kalman-Bucy filter is to estimate the state x as accurately as possible based solely on the measurements y_m. Put another way, the Kalman-Bucy filter attempts to regulate the norm of the state estimation error x_E to zero, where

$$x_E \equiv x - \hat{x}$$

and where the state estimate $\hat{x}$ shall be determined by a filter of the form (4.2). Mathematically, a cost function for this problem may thus be expressed as

$$\mathcal{J}_{KBF} \equiv E[\,||z_E||^2\,], \tag{4.9}$$

where $z_E \equiv x_E$ for notational convenience. (As Gaussian disturbances w_1 and w_2 continually drive this system, an integral on $t \in [0, \infty)$, as used to define $\mathcal{J}_{LQR}$, is not convergent for this problem, and the expectation value is the relevant measure.)

The mathematical statement of the present control problem, then, is the minimization of $\mathcal{J}_{KBF}$. This results in a "best possible" estimate of the state x. In order to arrive at

a form which is easily generalized in later sections, assume G_2 is nonsingular and define

$$B_1 \equiv \begin{pmatrix} G_1 & 0 \end{pmatrix} \qquad C_2 \equiv G_2^{-1} C \qquad D_{21} \equiv \begin{pmatrix} 0 & I \end{pmatrix}$$

and the vector of disturbances

$$w \equiv \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}.$$

Also, define new "observation" vectors y and $\hat{y}$ by a simple change of variables such that

$$y \equiv G_2^{-1}(y_m - D\,u) \qquad \hat{y} \equiv G_2^{-1}(\hat{y}_m - D\,u).$$

Note that this change of variables does not represent any real limitation, for whenever any flow measurement y_m is made in a physical implementation, the control u at that moment is also known, so the observation y is easily determined from the flow measurement y_m. With this change of variables, (4.8b) and (4.2b) may be expressed as

$$y = C_2\,x + D_{21}\,w \tag{4.10a}$$

$$\hat{y} = C_2\,\hat{x}. \tag{4.10b}$$

As we are developing the equations for an estimator, it is appropriate now to examine the equations for the state estimation error x_E and the ouput estimation error $y_E \equiv y - \hat{y}$. Subtracting (4.2a) from (4.8a) and (4.10b) from (4.10a) yields the system

$$\dot{x}_E = A\,x_E + B_1\,w + \hat{u} \tag{4.11a}$$

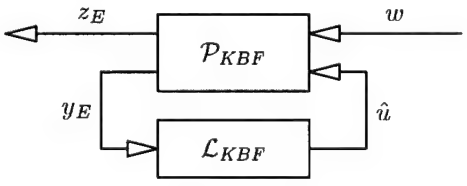
$$z_E = x_E \tag{4.11b}$$

$$y_E = C_2 x_E + D_{21} w. \tag{4.11c}$$

The Kalman-Bucy filter $\mathcal{L}_{KBF}$ is sought to relate the output injection term $\hat{u}$ to the output estimation error y_E. The extra term $\hat{u}$ is applied in these model equations to control the evolution of the state estimation error x_E such that the cost $\mathcal{J}_{KBF}(z_E)$ is minimized in the presence of Gaussian disturbances w. The important matrices of the system described by (4.11) may be summarized in the shorthand form

$$\mathcal{P}_{KBF} = \begin{array}{c} \\ \dot{x}_E \\ z_E \\ y_E \end{array} \begin{array}{c} \begin{array}{ccc} x_E & w & \hat{u} \end{array} \\ \left[\begin{array}{c|cc} A & B_1 & I \\ \hline I & 0 & 0 \\ C_2 & D_{21} & 0 \end{array}\right] \end{array}.$$

The flow of information is represented by the block diagram



where $\mathcal{P}_{KBF}$ is the flow system given by (4.11) and $\mathcal{L}_{KBF}$ is the Kalman-Bucy filter, which is still to be determined. This system accounts for Gaussian disturbances w and noisy observations y_E of the system, which are fed back through the filter $\mathcal{L}_{KBF}$ to produce the state estimate. The system output z_E may be used to monitor the performance of the system. Note the striking similarity of the structure of $\mathcal{P}_{KBF}$ to the structure of the conjugate transpose of $\mathcal{P}_{LQR}$. For this reason, these two problems are referred to as "duals", and their solutions are closely related.

Given this general setup, another Hamiltonian is defined such that

$$J_2 \equiv \begin{pmatrix} A^* & -C_2^* C_2 \\ -B_1 B_1^* & -A \end{pmatrix}. \tag{4.12a}$$

As shown in Doyle *et al.* (1989), the Hermetian positive-definite solution Y_2 to the algebraic Riccati equation defined by this Hamiltonian

$$A Y_2 + Y_2 A^* - Y_2 (C_2^* C_2) Y_2 + (B_1 B_1^*) = 0, \tag{4.12b}$$

denoted $Y_2 = \mathrm{Ric}(J_2)$, then yields the LTI estimator feedback matrix

$$L_2 = -Y_2 C_2^*. \tag{4.12c}$$

The LTI Kalman-Bucy filter $\mathcal{L}_{KBF}$ is then simply given by

$$\hat{u} = L_2 y_E,$$

and thus the complete state estimator is given by

$$\boxed{\dot{\hat{x}} = A \hat{x} + B_2 u - L_2 (y - C_2 \hat{x})} \tag{4.13}$$

This estimator minimizes $E[\,||x - \hat{x}||^2\,]$ in a system with Gaussian disturbances in the state equation and Gaussian noise in the measurements.

4.1.3. $\mathcal{H}_2$ *control (LQG = LQR + KBF)*

An estimator/controller of the form (4.2)—(4.3) for the complete system described by (4.8) with Gaussian disturbances may now be constructed. The objective of the control is to minimize

$$\mathcal{J}_2 \equiv E[\,||x||^2 + \ell^2 u^* u\,], \tag{4.14}$$

where $||\cdot||$ denotes the "2-norm" as defined in §2.5. Note that minimization of $\mathcal{J}_2$ is equivalent to minimization of the expectation value of $z^* z$, where

$$z \equiv \begin{pmatrix} Q^{1/2} x/\ell \\ u \end{pmatrix},$$

and Q is a diagonal matrix with diagonal entries $Q_{jj} = \pi/N$ as required by the definition of the norm in §2.4. As the control objective is the minimization of the expectation value of the square of a 2-norm, this type of estimator/controller is referred to as $\mathcal{H}_2$. As the state equation is linear, the cost quadratic, and the disturbances Gaussian, this type of estimator/controller is also referred to as linear quadratic Gaussian (LQG).

Combining the notation developed in the previous two sections

$$B_1 \equiv \begin{pmatrix} G_1 & 0 \end{pmatrix} \qquad C_1 \equiv \begin{pmatrix} Q^{1/2}/\ell \\ 0 \end{pmatrix} \qquad D_{12} \equiv \begin{pmatrix} 0 \\ I \end{pmatrix}$$
$$B_2 \equiv B \qquad C_2 \equiv G_2^{-1} C \qquad D_{21} \equiv \begin{pmatrix} 0 & I \end{pmatrix},$$

with the vector of disturbances w and the observation vectors y and $\hat{y}$ defined such that

$$w \equiv \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \qquad y \equiv G_2^{-1}(y_m - D u)$$
$$\hat{y} \equiv G_2^{-1}(\hat{y}_m - D u),$$

the system (4.8) and the control objective for the minimization of $\mathcal{J}_2$ take the form

$$\dot{x} = A\,x\ + B_1\,w + B_2\,u \tag{4.15a}$$
$$z = C_1 x + \qquad D_{12} u. \tag{4.15b}$$
$$y = C_2 x + D_{21} w. \tag{4.15c}$$

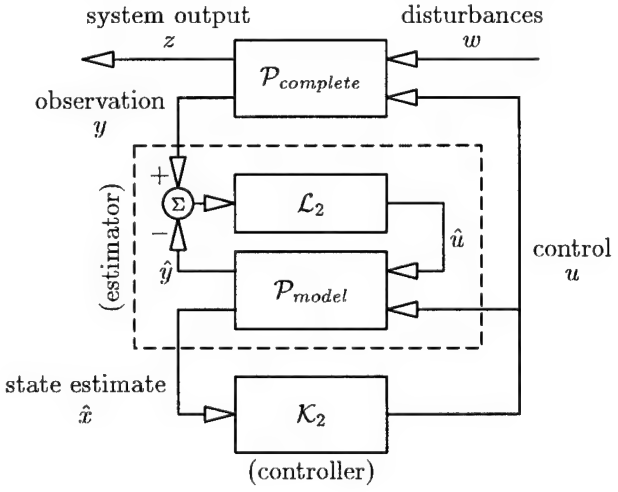
An $\mathcal{H}_2$ estimator/controller is sought to relate the observations y to the control u, which is applied to control the evolution of the state x such that the cost $\mathcal{J}_2(z)$ is minimized in the presence of Gaussian disturbances w. The important matrices of the complete system described by (4.15) may be summarized in the shorthand form

$$\mathcal{P}_{complete} = \begin{array}{c} \\ \dot{x} \\ z \\ y \end{array} \begin{array}{c} \begin{array}{ccc} x & w & u \end{array} \\ \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right] \end{array}.$$

Similarly, the important matrices of the model plant in an estimator of the form (4.2) may be represented in the shorthand form

$$\mathcal{P}_{model} = \begin{array}{c} \\ \dot{\hat{x}} \\ \hat{x} \\ \hat{y} \end{array} \begin{array}{c} \begin{array}{ccc} \hat{x} & \hat{u} & u \end{array} \\ \left[\begin{array}{c|cc} A & -I & B_2 \\ \hline I & 0 & 0 \\ C_2 & 0 & 0 \end{array}\right] \end{array}.$$

With these representations, the flow of information is represented by the block diagram



The plant, which is forced by external disturbances w, has an internal state x which cannot be observed. Instead, a few noisy observations y are made, and with these observations an estimate of the state $\hat{x}$ is determined. This state estimate is then fed back to through the controller to determine the control u to apply back on the plant in order to regulate x to zero. The system output z may be used to monitor the performance of the system.

The remarkable result from control theory (Lewis 1995) is that the $\mathcal{H}_2$ estimator/controller of the form illustrated in the above block diagram which minimizes $\mathcal{J}_2$ for this system is formed by simple combination of the optimal controller and the Kalman-Bucy filter such

that

$$\dot{\hat{x}} = A\,\hat{x} + B_2\,u - L_2\,(y - C_2\,\hat{x}) \tag{4.16a}$$

$$u = K_2\,\hat{x} \tag{4.16b}$$

where K_2 is given by (4.6)

$$K_2 = -B_2^*\,X_2 \qquad X_2 = \mathrm{Ric}\begin{pmatrix} A & -B_2\,B_2^* \\ -C_1^*\,C_1 & -A^* \end{pmatrix} \tag{4.16c}$$

and L_2 is given by (4.12)

$$L_2 = -Y_2\,C_2^* \qquad Y_2 = \mathrm{Ric}\begin{pmatrix} A^* & -C_2^*\,C_2 \\ -B_1\,B_1^* & -A \end{pmatrix}. \tag{4.16d}$$

Note the separation structure of this solution. The computation of K_2 does not depend upon the influence of the disturbances, which are accounted for in B_1 and C_2. The computation of L_2 does not depend upon the weightings in the cost function, which are accounted for in C_1, or the manner in which the control u affects the state, which is accounted for in B_2. In other words, the problem of control and the problem of state estimation are entirely decoupled.

Note also that, in order to arrive at the (relatively) simple control equations described by Doyle *et al.* (1989) and outlined in this section and the next, the matrices A, B_1, B_2, C_1, C_2, D_{12}, and D_{21} are assumed to satisfy eight required properties. The first four of these properties

$$(A, B_1) \text{ stabilizable} \qquad (C_1, A) \text{ detectable}$$

$$(A, B_2) \text{ stabilizable} \qquad (C_2, A) \text{ detectable}$$

are verified *a posteriori*, simply by examination of the results. (Note that the analysis of §3.3 indicates that there is only one slightly unstable mode for this system, and that this mode is both sensitive to the application of control and easily discerned by the measurements. Thus, we may presume, but not assert rigorously, that these four conditions will in fact be satisfied.) The matrices are constructed to satisfy the other four of these properties identically

$$D_{12}^*\,C_1 = 0 \qquad D_{12}^*\,D_{12} = I$$

$$B_1 D_{21}^* = 0 \qquad D_{21} D_{21}^* = I,$$

as may be verified directly with the definitions of these matrices.

4.2. $\mathcal{H}_\infty$ *control*

The $\mathcal{H}_\infty$ estimator/controller described in this section is very similar to the $\mathcal{H}_2$ estimator/controller described previously. Consideration is now given to disturbances, which we shall distinguish with a new variable χ, of the "worst" possible structure (as made precise below), rather than the Gaussian structure assumed in the $\mathcal{H}_2$ case. Considered in the frequency domain, the estimator/controllers developed in this section provide a system behaviour in which the maximum singular value of the closed-loop transfer function, also known as the "∞-norm", is less than some constant, which shall be referred to as γ. As this approach may be interpreted as bounding the ∞-norm of the transfer function from the disturbances to the performance measure, it is referred to as $\mathcal{H}_\infty$ control. For further details of the frequency-domain explanation of $\mathcal{H}_\infty$, the reader is referred to Doyle *et al.* (1989) and Zhou, Doyle, & Glover (1996).

The governing equations to be considered in this section are identical to (4.15):

$$\dot{x} = A\,x + B_1\,\chi + B_2\,u \tag{4.17a}$$

$$z = C_1 x + D_{12} u \tag{4.17b}$$

$$y = C_2 x + D_{21}\chi. \tag{4.17c}$$

As before, the G_1 and G_2 matrices used to define this system describe any covariance structure of the disturbances known or expected *a priori* (for instance, if one measurement is known to be noisier than another). These matrices are taken as identity matrices if no such structure is known in advance.

An $\mathcal{H}_\infty$ estimator/controller is sought to relate the observations y to the control u, which is applied to control the evolution of the state x such that the cost $\mathcal{J}_\infty(z)$ is minimized in the presence of some "worst case" disturbance χ. The flow of information is represented by a block diagram similar to that shown in §4.1.3.

Effectively, the cost function considered for $\mathcal{H}_\infty$ control is

$$\mathcal{J}_\infty \equiv E[x^*\,Q\,x + \ell^2\,u^* u - \gamma^2\,\chi^*\chi]. \tag{4.18}$$

A u is sought, through an estimator/controller of the form (4.2)—(4.3), to *minimize* $\mathcal{J}_\infty$, while *simultaneously* an external disturbance χ is sought to *maximize* $\mathcal{J}_\infty$. (In this manner, χ is the "worst possible" disturbance, as it is exactly that disturbance which increases the relevant cost function the most.) Thus, the $\mathcal{H}_\infty$ problem is a "min-max" problem. The term involving $-\gamma^2$ limits the magnitude of the unstructured disturbance in the maximization of $\mathcal{J}_\infty$ with respect to χ in a manner analogous to the term involving ℓ^2, which limits the magnitude of the control in the minimization of $\mathcal{J}_\infty$ with respect to u.

The result (Doyle *et al.* 1989) is that an $\mathcal{H}_\infty$ estimator/controller of the form (4.2)—(4.3) which minimizes $\mathcal{J}_\infty$ in the presence of some component of the worst case unstructured disturbance χ for this system is given by

$$\dot{\hat{x}} = A\,\hat{x} + B_2\,u - L_\infty\,(y - C_2\,\hat{x}) \tag{4.19a}$$

$$u = K_\infty\,\hat{x} \tag{4.19b}$$

where K_∞ is given by

$$K_\infty = -B_2^*\,X_\infty \qquad X_\infty = \mathrm{Ric}\begin{pmatrix} A & \gamma^{-2}\,B_1\,B_1^* - B_2\,B_2^* \\ -C_1^*\,C_1 & -A^* \end{pmatrix} \tag{4.19c}$$

and L_∞ is given by

$$L_\infty = -Y_\infty\,C_2^* \qquad Y_\infty = \mathrm{Ric}\begin{pmatrix} A^* & \gamma^{-2}\,C_1^*\,C_1 - C_2^*\,C_2 \\ -B_1\,B_1^* & -A \end{pmatrix}. \tag{4.19d}$$

Note first that, in the $\gamma \to \infty$ limit, the $\mathcal{H}_2$ estimator/controller is recovered, so the set of two Riccati equations in (4.19) describes both the $\mathcal{H}_2$ (optimal control + Kalman-Bucy filter) and the $\mathcal{H}_\infty$ problems.

It may also be shown that, as the upper-right blocks of the Hamiltonians may not be negative definite, a solution to these Riccati problems exists only for sufficiently large γ; the smallest $\gamma = \gamma_0$ for which a solution to these equations exists may be found by trial and error (Doyle *et al.* 1989). An $\mathcal{H}_\infty$ estimator/controller for $\gamma > \gamma_0$ is referred to as suboptimal.

4.3. *Comparison of $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control equations*

Most of the robustness problems associated with $\mathcal{H}_2$ stem from the state estimation. Optimal (LQR) controllers themselves, provided with full state information, generally have excellent performance and robustness properties (Dailey *et al.* 1990). Recall from §4.1.3 that the problems of control and state estimation in the $\mathcal{H}_2$ formulation are decoupled.

An important observation of §4.2 is that the problems of control and state estimation in the $\mathcal{H}_\infty$ formulation are coupled. Specifically, the computation of K_∞ depends on the expected covariance of the state disturbances, which are accounted for in B_1, and the computation of L_∞ depends on the weightings in the cost function, which are accounted for in C_1. This is one of the essential features of $\mathcal{H}_\infty$ control.

By taking into account the expected covariance of the state disturbances, reflected in B_1, when determining the state feedback matrix K_∞, the components of $\hat{x}$ corresponding to the components of x that are expected to have the smallest forcing by external disturbances are weighted least in the feedback control relationship $u = K_\infty \hat{x}$.

Similarly, by taking into account the weightings in the cost function, reflected in C_1, when determining the estimator feedback matrix L_∞, the components of $\hat{x}$ corresponding to the components of x that are least important in the computation of $\mathcal{J}_\infty$ are forced with the smallest corrections by the output injection term $L_\infty (y - \hat{y})$ in the equation for the estimator.

By applying strong control only on those components of $\hat{x}$ significantly excited by external disturbances, and by applying strong estimator corrections only to those components of $\hat{x}$ important in the computation of the cost function, $\mathcal{H}_\infty$ feedback gains for components of the system not relevant to the control problem are reduced from those in the $\mathcal{H}_2$ case. With such feedback gains reduced, the stability properties of $\mathcal{H}_\infty$ estimator/controllers in the presence of state disturbances and measurement noise may be expected to be better than their $\mathcal{H}_2$ counterparts, at the cost of a (hopefully, small) degradation of performance in terms of the 2-norm of the output z for the undisturbed system.

4.4. *Numerical method*

Standard numerical techniques are now applied to all aspects of this problem. In order to simplify both the theory to be presented and the numerical algorithm to be coded, no further manipulation of the equations is used beyond the matrix representations (4.17) and (4.19). It was observed that the minimal realization approach (Kailath 1980) is well suited to reduce the computation time necessary to determine effective control algorithms by the present approach; however, such an approach was not found to be necessary in the present case.

The algebraic Riccati equations are solved using the method of Laub (1991), which involves a Schur factorization. This is found to be a stable numerical algorithm for all cases tested. The implementation of Laub's method is written in Fortran-90 and follows closely the algorithm used by the Matlab function `are.m` (Grace *et al.* 1992). A Lyapunov solver, modelled after the Matlab function `lyap.m`, is also used to compute the system Gramians.

Two LAPACK routines (Anderson *et al.* 1995), `zgeev.f` and `zgees.f`, are used to compute eigenvalues/eigenvectors and Schur factorizations. These routines are compiled in quad precision (128 bits per real number) to ensure sufficient numerical precision in the eigenvalue computation. All computations are carried out with $N = 140$ to ensure good resolution of all significant eigenmodes. The eigenvalues of A match all those tabulated by Orszag (1971) to all eight decimal places, as shown in table 1, indicating that this numerical method is sufficiently accurate.

5. Performance of controlled systems

We now examine the behaviour of the "closed-loop" systems obtained by application of the above controllers and estimators to the "nominal" (*i.e.* no disturbances) channel flow stability problem. In other words, we examine the behaviour of the flow and the estimator/controllers operating together as a single dynamical system. By looking at "root locus" plots which map the movement of the eigenvalues of these systems in the complex plane with respect to the relevant parameters, this behaviour is well quantified, as long as all perturbations remain small enough that the linearity assumption remains valid. We shall also examine the control and observation sensitivities defined in §3.2 for two special cases in order to better understand the fundamental limitations of controllers and estimators applied to the present system.

5.1. $\mathcal{H}_2$ *control*

5.1.1. *Optimal control (LQR)*

In order to investigate the controllability of the closed-loop eigenmodes when all modes are observable, consider the system described in §4.1.1. With $r = 0$ and examining only the equations for $\dot{x}$ and $\hat{x}$, the plant is given (in the shorthand notation used in §4) by

$$\mathcal{P}_{LQR} = \begin{array}{c} \\ \dot{x} \\ \hat{x} \end{array} \begin{array}{c} \begin{array}{cc} x & u \end{array} \\ \left[\begin{array}{c|c} A & B_2 \\ \hline I & 0 \end{array}\right] \end{array}$$

with the control now given by

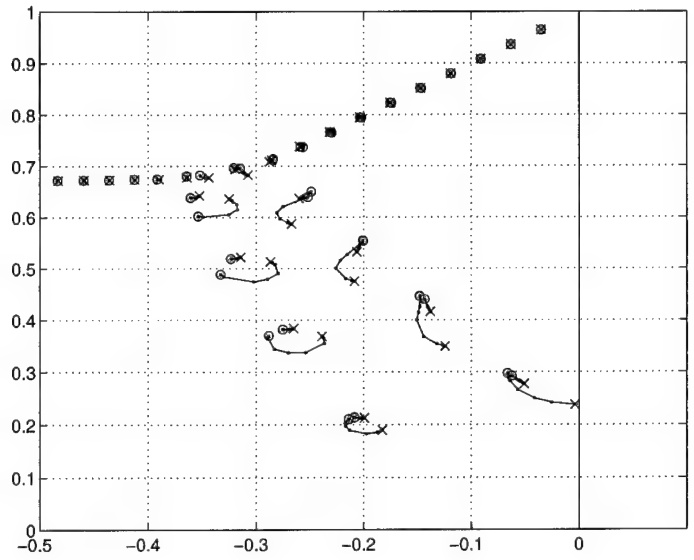
$$u = K_2\,\hat{x} + u',$$

where an additional control term u' has been added to study the sensitivity of the closed-loop system to further modification of the control. Putting the plant and the controller together, the closed-loop system may be represented by

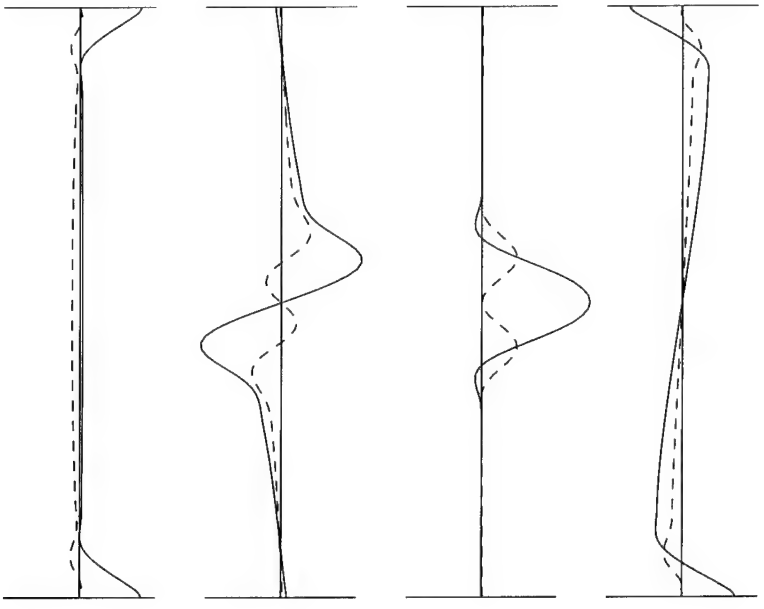
$$\mathcal{P}_{LQR(closed\ loop)} = \begin{array}{c} \\ \dot{x} \\ \hat{x} \end{array} \begin{array}{c} \begin{array}{cc} x & u' \end{array} \\ \left[\begin{array}{c|c} A + B_2\,K_2 & B_2 \\ \hline I & 0 \end{array}\right] \end{array}. \tag{5.1}$$

The eigenmodes of $A_{K_2} \equiv A + B_2\,K_2$ describe the dynamics of the closed-loop system for the unmodified control rule ($u' = 0$). Figure 2a shows the movement of these eigenvalues with respect to the free parameter of the control problem, ℓ, used to determine K_2. The eigenvalues for $\ell \to \infty$ are observed to be very near those of the uncontrolled system A in figure 1a, with the previously unstable mode moved just to the left of the imaginary axis. The eigenvalues generally move to the left as ℓ is decreased. Figure 2b shows the shape of the first four eigenmodes of the closed-loop system. Comparing figure 2b with figure 1b, it is seen that the control modifies most those eigenmodes with significant variations near the wall.

The sensitivity of the eigenmodes of the system (5.1) to modification of the control rule may be quantified by performing the analysis of §3.2.1, replacing the eigenmodes of A by the eigenmodes of A_{K_2}. The result of this analysis for small ℓ is shown in table 2. This table shows that, in the $\ell \to 0$ limit, the system matrix is modified to the point that the eigenmodes are no longer sensitive to further modification of the control. In other words, all the controllable dynamics of the system have been modified by K_2 and are accounted for in the closed loop system in this limit. This is one demonstration that the optimal controller extracts the best possible performance from a given (full-information) system.



(a) Root locus of least stable eigenvalues of A_{K_2} as a function of the free parameter of the $\mathcal{H}_2$ controller, ℓ. The eigenvalues for $\ell \to \infty$ are marked with an (x).



(b) Eigenvectors of A_{K_2}, with $\ell = 10^{-4}$, corresponding to (left to right): $j = 1$, $j = 3$, $j = 4$, and $j = 5$. Corresponding eigenvalues are reported in table 2.

FIGURE 2. Least stable eigenmodes of A_{K_2}.

j	λ_j	f_j
3	$-0.03513233 - 0.96462128i$	0.000000029
4	$-0.03518652 - 0.96464261i$	0.000000001
5	$\mathbf{-0.06255259 - 0.29262711i}$	**0.000001101**
6	$-0.06310358 - 0.93629329i$	0.000000070
7	$-0.06325089 - 0.93635257i$	0.000000003
1	$\mathbf{-0.06644730 - 0.29721403i}$	**0.000001116**
8	$-0.09102975 - 0.90793951i$	0.000000129
9	$-0.09130964 - 0.90805917i$	0.000000008
10	$-0.11890731 - 0.87955083i$	0.000000226
11	$-0.11936036 - 0.87976246i$	0.000000020
12	$\mathbf{-0.14335180 - 0.43962023i}$	**0.000002303**
14	$-0.14673294 - 0.85111508i$	0.000000414
15	$-0.14739907 - 0.85146161i$	0.000000045
13	$\mathbf{-0.14803996 - 0.44586838i}$	**0.000003081**
16	$-0.17450455 - 0.82261690i$	0.000000842

TABLE 2. Least stable eigenmodes of the closed-loop system A_{K_2} and their sensitivity to control for the optimal controller in the cheap control limit ($\ell = 10^{-4}$). The numbering of the eigenvalues shown is the same as the numbering of the eigenvalues of table 1 to which they are connected by the root locus of figure 2. Note that the control in this limit drives all eigenmodes to positions at which they are insensitive to further modifications of the control, as illustrated by the large reductions in f_j. Note also that those eigenmodes with the largest values of f_j in table 1 (specifically, those in the lower branch) have moved the most.

5.1.2. *Kalman-Bucy filter (KBF)*

The estimator itself has its own set of dynamics. These dynamics are captured by the equations for the state estimator error, as described in §4.1.2. We now make use of this system in order to investigate the observability of closed-loop eigenmodes when all modes are controllable. With $w = 0$ and examining only the equations for $\dot{x}_E$ and y_E, this plant is given by

$$\mathcal{P}_{KBF} = \begin{array}{c} \\ \dot{x}_E \\ y_E \end{array} \begin{array}{c} \begin{array}{cc} x_E & \hat{u} \end{array} \\ \left[\begin{array}{c|c} A & I \\ \hline C_2 & 0 \end{array}\right] \end{array}$$

with the output injection now given by

$$\hat{u} = L_2\, y_E + \hat{u}',$$

where an additional output injection term $\hat{u}'$ has been added to study the sensitivity of the closed-loop system to further modification of the output injection rule. Putting the plant and the estimator together, the closed-loop system may be represented by

$$\mathcal{P}_{KBF\,(closed\ loop)} = \begin{array}{c} \\ \dot{x}_E \\ y_E \end{array} \begin{array}{c} \begin{array}{cc} x_E & \hat{u}' \end{array} \\ \left[\begin{array}{c|c} A + L_2\, C_2 & I \\ \hline C_2 & 0 \end{array}\right] \end{array}. \tag{5.2}$$

The eigenmodes of $A_{L_2} \equiv A + L_2\, C_2$ describe the dynamics of the closed-loop system for the unmodified output injection rule ($\hat{u}' = 0$). Figure 3 shows the movement of these eigenvalues with respect to the free parameters of the estimator problem. (This is done by assuming that the matrices describing the covariance of the disturbances have the simple form $G_1 = g_1\, I$ and $G_2 = g_2\, I$, where g_1 and g_2 are real scalars.) The eigenvalues for $g_1 = g_2 \to 0$ are very near those of the uncontrolled system A in figure 1a, with the

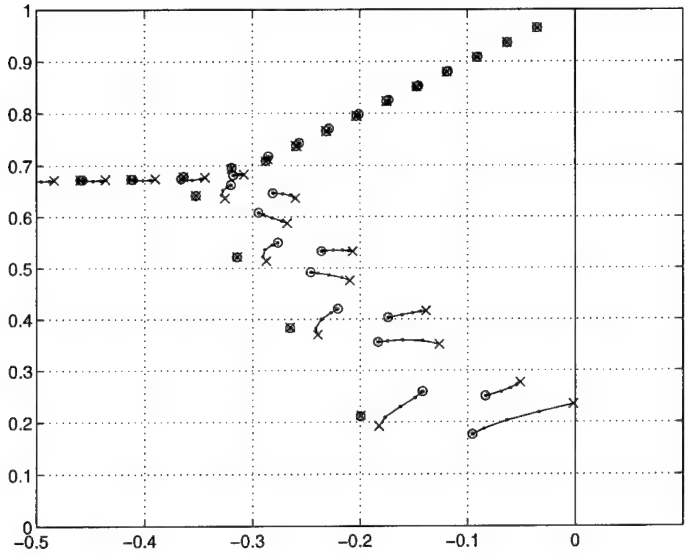


FIGURE 3. Root locus of least stable eigenvalues of A_{L_2} as a function of the free parameters of the $\mathcal{H}_2$ estimator, g_1 and g_2 (note that we take $g_1 = g_2$ for the purpose of drawing the root locus). The eigenvalues for $g_1 = g_2 \to 0$ are marked with an (x).

j	λ_j	g_j
3	$-0.03505745 - 0.96474093i$	0.000000568
4	$-0.03518656 - 0.96464253i$	0.000000004
6	$-0.06287931 - 0.93668086i$	0.000000644
7	$-0.06325136 - 0.93635193i$	0.000000008
5	$\mathbf{-0.08362450 - 0.25066856i}$	**0.000002858**
8	$-0.09059621 - 0.90874817i$	0.000000673
9	$-0.09131196 - 0.90805689i$	0.000000011
1	$\mathbf{-0.09565183 - 0.17658643i}$	**0.000000094**
10	$-0.11823779 - 0.88095122i$	0.000000646
11	$-0.11936807 - 0.87975709i$	0.000000014
12	$\mathbf{-0.14209547 - 0.25910275i}$	**0.000000130**
14	$-0.14584717 - 0.85329567i$	0.000000549
15	$-0.14741926 - 0.85145223i$	0.000000014
16	$-0.17347707 - 0.82577419i$	0.000000399
13	$\mathbf{-0.17418920 - 0.40314656i}$	**0.000002002**

TABLE 3. Least stable eigenmodes of the closed-loop system A_{L_2} and their sensitivity to observation for the Kalman-Bucy filter in the large disturbance limit ($g_1 = g_2 = 10^2$). The numbering of the eigenvalues shown is the same as the numbering of the eigenvalues of table 1 to which they are connected by the root locus of figure 1. Note that the estimator in this limit modifies all eigenmodes until the measurements are no longer sensitive to them, as illustrated by the large reductions in g_j. Note also that those eigenmodes with the largest values of g_j in table 1 (specifically, those in the lower branch) have moved the most.

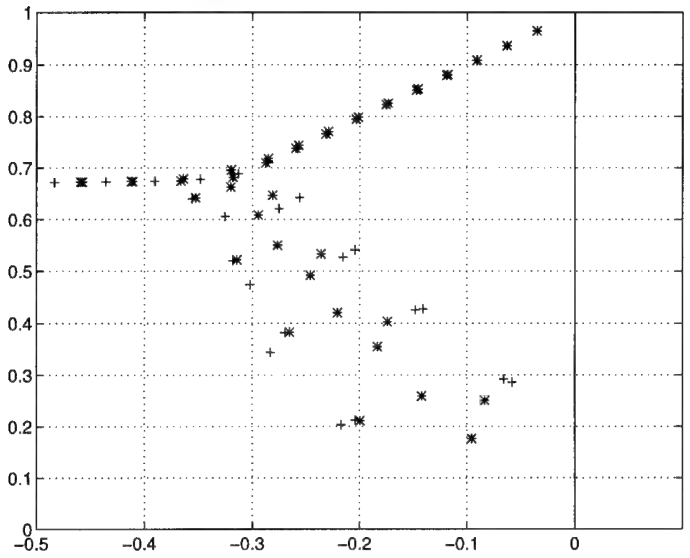


FIGURE 4. Least stable eigenvalues of the composite closed-loop system with the $\mathcal{H}_2$ estimator/controller, taking $\ell = g_1 = g_2 = 1$. Note that the eigenvalues are simply the eigenvalues of the closed loop controller (+) together with those of the closed loop estimator (*).

previously unstable mode moved just to the left of the imaginary axis. The eigenvalues generally move to the left as g_1 and g_2 are increased.

The sensitivity of measurements y_E to the the eigenmodes of the system (5.2) may be quantified by performing the analysis of §3.2.2, replacing the eigenmodes of A by the eigenmodes of A_{L_2}. The result of this analysis for large $g_1 = g_2$ is shown in table 3. This table shows that, in the $g_1 = g_2 \to \infty$ limit, the system matrix is modified to the point that the measurements are no longer sensitive to the eigenmodes of the closed-loop system. In other words, all the measurable dynamics of the system have been extracted by L_2 and are accounted for in the closed loop system in this limit. This is one demonstration that the Kalman-Bucy filter extracts the best possible state estimate from a given (fully-controllable) state estimator.

5.1.3. *$\mathcal{H}_2$ control (LQG = LQR + KBF)*

It was mentioned in §4.1.3 that the estimator/controller which minimized the relevant cost functional ($\mathcal{J}_2$) in the presence of Gaussian disturbances could be found by considering the controller and estimator problems separately. In this section, it is shown that the closed-loop performance of a system of the form (4.15) (without disturbances)

$$\begin{aligned} \dot{x} &= A\,x + B_2\,u \\ y &= C_2 x \end{aligned}$$

combined with an estimator/controller of the form (4.16)

$$\begin{aligned} \dot{\hat{x}} &= A\,\hat{x} + B_2\,u - L_2\,(y - C_2\,\hat{x}) \\ u &= K_2\,\hat{x} \end{aligned}$$

may also be evaluated by considering the estimator and controller problems separately. To accomplish this, simply combine the above equations into the closed-loop composite

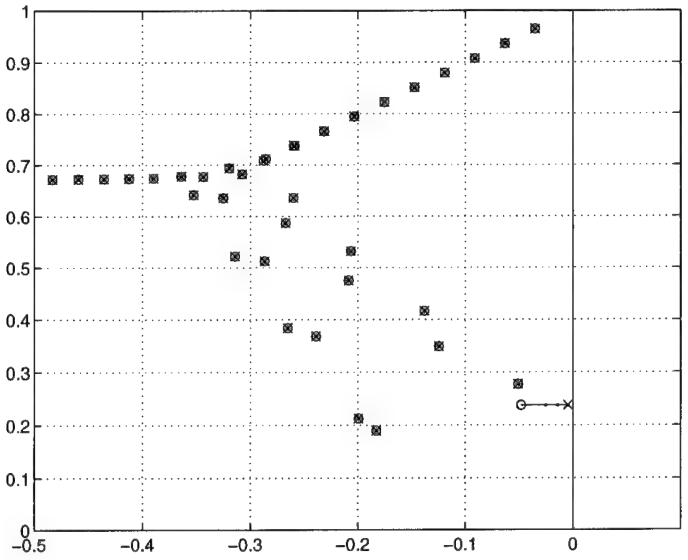


FIGURE 5. Root locus of least stable eigenvalues of the $\mathcal{H}_\infty$ controller versus γ, taking $\ell = 100$, $g_1 = g_2 = 0.001$. The result with $\gamma \to \infty$, marked with the (x), gives the corresponding $\mathcal{H}_2$ controller. Note that the $\mathcal{H}_\infty$ controller modifies only the least stable eigenmode of this $\mathcal{H}_2$ result, without expending any extra control effort to control those eigenmodes not associated with the maximally unstable component of the system. Note also that $\gamma = \gamma_0$, marked with the (o), is reached by reducing γ until the least stable eigenvalue corresponds to one of the uncontrollable eigenmodes in the upper branch, which cannot be moved further left; in the present case, this corresponds to a numerical value of $\gamma_0 = 0.26$.

system

$$\begin{pmatrix} \dot{x} \\ \dot{\hat{x}} \end{pmatrix} = \begin{pmatrix} A & B_2\,K_2 \\ -L_2\,C_2 & A + B_2\,K_2 + L_2\,C_2 \end{pmatrix} \begin{pmatrix} x \\ \hat{x} \end{pmatrix}.$$

Gaussian elimination, first on the rows and then on the columns, reveals that the eigenvalues of this system are the same as the eigenvalues of the system

$$\begin{pmatrix} A + B_2\,K_2 & B_2\,K_2 \\ 0 & A + L_2\,C_2 \end{pmatrix}.$$

In other words, the eigenvalues of the closed-loop composite system for the $\mathcal{H}_2$ problem are simply the union of the eigenvalues of the controlled system $A_{K_2} = A + B_2\,K_2$ and the eigenvalues of the estimated system $A_{L_2} = A + L_2\,C_2$ discussed in the previous two sections and illustrated in figure 4.

5.2. $\mathcal{H}_\infty$ *control*

As with the $\mathcal{H}_2$ estimator/controller, the performance of the closed loop composite system with the $\mathcal{H}_\infty$ estimator/controller

$$\begin{pmatrix} \dot{x} \\ \dot{\hat{x}} \end{pmatrix} = \begin{pmatrix} A & B_2\,K_\infty \\ -L_\infty\,C_2 & A + B_2\,K_\infty + L_\infty\,C_2 \end{pmatrix} \begin{pmatrix} x \\ \hat{x} \end{pmatrix}.$$

may be evaluated by considering the performance of the controlled system $A_{K_\infty} = A + B_2\,K_\infty$ and the performance of the estimated system $A_{L_\infty} = A + L_\infty\,C_2$ separately. The root locus of the eigenvalues of A_{K_∞} are plotted with respect to the parameter γ

of the H_∞ problem in figure 5, clearly illustrating the tendency of $\mathcal{H}_\infty$ controllers to modify only the least stable components of the system, as opposed to the $\mathcal{H}_2$ controller of figure 2, which modifies all controllable modes of the system.

6. Conclusions

Optimal and robust control theories have been successfully applied to the Orr-Sommerfeld equation. Given control on the wall-normal component of boundary velocity only, the flow system is shown to be stabilizable but not controllable. Given measurements of wall skin-friction only, the flow system is shown to be detectable but not observable. It is shown that $\mathcal{H}_2$ controllers/estimators modify all of the controllable/observable modes of the system. In contrast, the $\mathcal{H}_\infty$ controllers modify the corresponding $\mathcal{H}_2$ controllers only in the most unstable component, as $\mathcal{H}_\infty$ targets a bound only on the maximum value of the transfer function.

In the $\ell \to 0$ limit of the $\mathcal{H}_2$ controller, corresponding to cheap control and thus large values of u, all eigenmodes of the closed-loop controlled system are shown to be modified to points at which they are no longer sensitive to further modifications of the control. Similarly, in the $g_1 = g_2 \to \infty$ limit of the $\mathcal{H}_2$ estimator, accounting for large disturbances on both the state and the measurements, all eigenmodes of the closed-loop system for the estimator error are shown to be modified to points at which they are not discernible by flow measurements.

These results indicate that $\mathcal{H}_2$ controllers and estimators are optimal for their desired purposes, but may contain large feedback gains. On the other hand, $\mathcal{H}_\infty$ controllers only target the least stable components of the system, and thus have smaller feedback gains while still achieving the same worst case performance for the nominal plant. Such reduced feedback gains generally result in improved robustness to inaccuracies in the system model.

REFERENCES

Abergel, F. & Temam, R. 1990 On some control problems in fluid mechanics *Theor. and Comp. Fluid Dynamics* **1**, 303-325.

Anderson, E., Bai, Z., Bischof, C., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A., Ostrouchov, S., and Sorensen, D. 1995 *LAPACK Users' Guide.* SIAM.

Bridges, T.J., & Morris, P.J. 1984 Differential eigenvalue problems in which the parameter appears nonlinearly *J. Comput. Phys.* **55**, 437-460.

Butler, K.M. & Farrell, B.F. 1992 Three-dimensional optimal perturbations in viscous shear flows *Phys. Fluids A* **4**, 8, 1637-1650.

Canuto, C., Hussaini, M.Y., Quarteroni, A., & Zang, T.A. 1988 *Spectral Methods in Fluid Dynamics.* Springer.

Dailey, R.L., Doyle, J.C., Stein, G., Banda, S.S., & Yeh, H.H. 1990 *Lecture Notes for the Workshop on $\mathcal{H}_\infty$ and μ Methods for Robust Control* Presented at 1990 American Control Conference, May 21-22, San Diego.

Doyle, J.C., Glover, K., Khargonekar, P.P., & Francis, B.A. 1989 State-Space Solutions to Standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Control Problems *IEEE Trans. Auto. Control* **34**, 8, 831-847.

Franklin, G.F., Powell, J.D., and Emami-Naeini, A. 1991 *Feedback Control of Dynamic Systems.* Addison-Wesley.

Grace, A., Laub, A.J., Little, J.N., & Thompson, C.M. 1992 *Control System Toolbox User's Guide.* The MathWorks, Inc.

Joshi, S.S., Speyer, J.L., & Kim, J. 1995 Modelling and control of two dimensional Poiseuille flow, *34th IEEE Conference on Decision and Control*, 921-927.

JOSHI, S.S., SPEYER, J.L., & KIM, J. 1996 A systems theory approach to the feedback stabilization of infinitesimal and finite-amplitude disturbances in plane Poiseuille flow, accepted for publication in *J. Fluid Mech.*

JOSLIN, R.D., GUNZBURGER, M.D., NICOLAIDES, R.A., ERLEBACHER, G., & HUSSAINI, M.Y. 1995 A self-contained, automated methodology for optimal flow control validated for transition delay *ICASE Report No 95-64*, NASA Langley Research Center, Hampton, VA.

KAILATH, T. 1980 *Linear Systems.* Prentice-Hall.

KHORRAMI, M.R., MALIK, M.R., & ASH R.L. 1989 Application of spectral collocation technique to the stability of swirling flows *J. Comput. Phys.* **81**, 206-229.

KWAKERNAAK, H., & SIVAN, R. 1972 *Linear Optimal Control Systems.* Wiley.

LAUB, A.J. 1991 Invariant subspace methods for the numerical solution of Riccati equations. In *The Riccati Equation* (ed. Bittaini, Laub, & Willems) 163-196. Springer.

LEWIS, F.L. 1995 *Optimal Control.* Wiley.

MOIN, P., & BEWLEY, T.R. 1995 Application of control theory to turbulence *Twelfth Australasian Fluid Mechanics Conference*, Dec. 10-15, Sydney, 109-117.

ORSZAG, S.A. 1971 Accurate solution of the Orr-Sommerfeld stability equation *J. Fluid Mech.* **50**, 4, 689-703.

ORSZAG, S.A., & PATERA, A.T. 1983 Secondary instability of wall-bounded shear flows *J. Fluid Mech.* **128**, 347-385.

ZHOU, K., DOYLE, J.C., & GLOVER, K. 1996 *Robust and Optimal Control.* Prentice-Hall.

PART B.

Optimal control of turbulence

Optimal control theory is used to determine controls that effectively reduce the drag of a turbulent flow in a plane channel. Wall transpiration (unsteady blowing/suction) with zero net mass flux is used as the control. The technique described is unique from the standpoint that it is mathematically based solely on the control objective, the equations governing the fluid flow, and instantaneous observations of the flow, without the *ad hoc* procedures normally used to accomplish control of complex nonlinear phenomena such as turbulence. Drag reduction of over 50 percent is obtained using an optimal controller in a direct numerical simulation of a turbulent channel flow at $Re_\tau = 180$, which far exceeds what has been obtained to date via adaptive and intuition-based control rules in similar flows.

The algorithm used is computationally intensive and requires full flowfield information, and therefore can not be implemented in a laboratory. However, these calculations allow us to quantify the best possible system performance given a certain class of flow actuation and qualitatively identify how optimized controllers interact with the coherent structures of the turbulence. In so doing, an important step is made in the progression towards practical and efficient turbulence control strategies based on optimal control techniques.

1. Introduction

The recent development of the technology necessary to produce micro-scale mechancanical devices†, commonly referred to as Microfabricated Electro-Mechanical Systems (MEMS), has prompted researchers to revisit questions heretofore thought to be purely academic. The present question is exactly of this nature: assuming that individual small-scale turbulent fluctuations may somehow be measured and that concomitant small-scale forcing of the turbulent fluid may somehow be attained, how much do practical engineering designs stand to benefit, where should the control be applied, and what control algorithms are most effective? The present work attempts to cast these questions in a rigorous framework, present a mathematical approach for their solution, and demonstrate an effective implementation of the control algorithm in a fully-developed turbulent flow.

We first briefly summarize recent approaches to determine effective turbulence control algorithms, categorizing these approaches to the feedback control problem by examining their mathematical dependence on the equations governing the system. This discussion puts the present approach in context with other techniques currently under investigation. For a more thorough discussion along this line, see Moin & Bewley (1994).

1.1. *Adaptive networks*

The first class of schemes which may be proposed to achieve small scale flow control actually makes no explicit reference to the dynamics known to take place in the flow or the Navier-Stokes equations known to govern these dynamics. Instead, a "reasonable" network is fashioned which takes as input those measureable flow quantities assumed to

† For a recent review of this subject as it applies to fluid mechanics, see McMichael (1996).

be most relevant to the control problem and produces as output the requisite control velocity. The coefficients of this network are then "trained" by applying the control network to the flow and gradually adjusting the coefficients in a heuristic manner based on the resulting evolution of the flow.‡

As an example of one adaptive approach, an adaptive inverse technique has been applied by Lee *et al.* (1996) to a turbulent channel flow at $Re_\tau = 100$, providing approximately 20 percent drag reduction. This approach first develops an approximate "inverse" model between measurable flow quantities (as input) and the control forcing (as output) with an adaptive technique. Each iteration of the adaptation for this inverse model consists of three steps: 1) computing the error of the model output with respect to the desired model output (the actual control forcing used), 2) determining the influence of the various weights in the model on this error, then 3) updating all the weights in the model a small amount in a manner that reduces the error. When applied to the nonlinear adaptive networks commonly used for this purpose, known as "neural networks", this is commonly referred to as "back-propagation" of the error. Once the approximate inverse model between the flow measurements and the control converges, the inverse model is used to compute a control which will drive the flow measurements to some desired state. In the case of Lee *et al.* (1996), the desired state is chosen to be a state with reduced spanwise drag fluctuations, and the inverse model is continually trained as the flow system evolves in time. Such on-line training of the controller helps to provide "robustness" to possible changes in the dynamics of the flow system to be controlled.

1.2. *Intuition-based approaches*

In situations in which the dominant physics is well understood, judgment can guide an engineer to design effective control schemes. Success is limited, however, by the engineer's understanding of the physical processes involved; in the case of turbulence, our understanding is still limited despite several decades of intense research.

An active cancellation scheme was used by Choi, Moin, & Kim (1994), to reduce the drag in a fully-developed turbulent flow by mitigating the effect of the near-wall vortices. By opposing the near-wall motions of the fluid, which are caused by the near-wall vortices, with an opposing wall control, the high shear region was lifted away from the wall. A direct numerical simulation of this scheme applied to turbulent channel flow at $Re_\tau = 100$ demonstrated 23 percent drag reduction when the control was chosen to oppose the vertical motion at $y^+ = 10$.

Sensing the instantaneous normal velocity at $y^+ = 10$ is, of course, very difficult. From a practical standpoint, it is highly desirable to confine both sensing and actuation to the wall. Thus, Choi, Moin, & Kim (1994) computed the correlation of quantities measurable at the wall with the normal velocity above the wall. Surprisingly, the wall pressure did not exhibit a high correlation with the normal velocity. Using a Taylor series expansion and the equation of continuity, they obtained an expression relating the normal velocity at a point near the wall to the instantaneous wall shear. However, using this expression to estimate the normal velocity away from the wall resulted in only a 6 percent drag reduction. This is comparable to the drag reduction that can be achieved with simpler, passive means such as riblets.

‡ Note that the network used for this purpose can take any of several linear or nonlinear forms. Hertz, Krogh, & Palmer (1991) contains a survey of these networks and outlines strategies for their training.

1.3. *Dynamical systems*

The tools of dynamical systems theory have proven useful in analyzing and interpreting turbulence dynamics (Aubry *et al.* 1988). Due to their large range of spatial and temporal scales, turbulent flows are known to have relatively high dimensions in this framework even at fairly low Reynolds numbers, which makes analysis of these systems quite difficult (Keefe, Moin, & Kim 1992). However, there has been some instructive work in representing the dynamics of coherent structures in wall-bounded flows with systems of much lower dimension using the proper orthogonal decomposition (POD) method (Berkooz, Holmes, & Lumley 1993). This method provides a (numerically determined) set of eigenfunctions which are particularly efficient in representing second order turbulence statistics with a small number of modes (perhaps as few as 10 to 20) in the cross-flow plane.

In the dynamical systems framework, the movement of the near-wall longitudinal vortices when observed in a cross-flow plane may be represented locally as the orbiting of a low dimensional state around several unstable fixed points; the passage of one set of coherent structures leads to a rapid jump in the state to a different unstable orbit, or to a different distribution of near-wall longitudinal vortices in the cross-flow plane. Such a rapid transition between critical points, followed by a quiescent period in which the flow pattern remains largely unchanged, is referred to as a heteroclinic cycle.

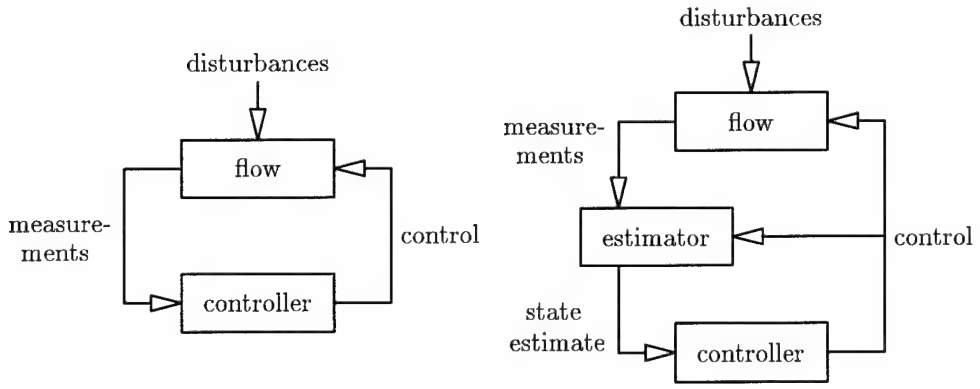
Coller, Holmes, & Lumley (1994a,b) consider the control of an interesting model problem governed by a simple two-component equation with similar dynamics to this model of near-wall longitudinal vortices (*i.e.* attracting heteroclinic cycles) subject to random excitation to account for unmodelled disturbances. They develop and demonstrate a strategy which delays heteroclinic transitions in this model as long as possible by sensing when the state is near an unstable fixed point and maintaining it there with feedback control for as long as possible. Once the state diverges from this fixed point, presumably due to the unmodelled dynamics of the flow, control is turned off until the state approaches the neighborhood of another unstable fixed point.

1.4. *Rigorous optimization of practical control algorithms—a preview*

The present work is a first step towards developing practical controllers of the two types illustrated in figure 1. The (initially undetermined) coefficients present in both configurations may be optimized rigorously with approaches based on the adjoint analysis developed in this paper, and are still under development (Bewley, Moin, & Temam 1996).

In the output feedback configuration, the flow is controlled using computationally inexpensive direct feedback from instantaneous flow measurements. The structure of the controller may be nonlinear and may incorporate a finite impluse response (FIR) filter to account for information from past measurements in the control rule.

In the estimator/controller configuration, a time-evolving estimate of the flow state near the wall is first developed, effectively accumulating the information reflected by the stream of measurements from a few noisy sensors. The flow is then controlled with a (possibly nonlinear) control rule based on this flow estimate. The estimation problem and the control problem become linked when they are optimized in the presence of a small component of "worst case" noise which maximally aggrevates the coupled system. Such an approach is well developed for linear problems, and is referred to as $\mathcal{H}_\infty$ control; Doyle *et al.* (1989) presents a compact form of this approach which makes it straightforward to apply to linear problems, as illustrated in Bewley, Agarwal, & Liu (1996) for the control of the linear stages of transition. Methods to extend this "robust" approach to nonlinear problems are still under analysis.



a) Output feedback configuration *b) Estimator/controller configuaration*

FIGURE 1. Examples of practical control configurations based on a few noisy flow measurements. Both closed-loop configurations may be optimized with techniques based on the approach developed in this paper, and are still under development.

2. Governing equations

The flow problem we will consider is fully-developed turbulent channel flow with no-slip walls and wall-normal velocity boundary conditions Φ applied as the control. Though this is an idealized geometry, it will give insight into the turbulent behavior which can later be exploited in more practical configurations, such as the control of a spatially developing boundary layer with discrete wall-mounted actuators. The present problem is governed by the unsteady, incompressible Navier-Stokes equation and the continuity equation inside the domain Ω and wall-normal velocity boundary conditions on the walls $\partial\Omega_1$. On the remainder of the boundary of the three-dimensional volume Ω, denoted $\partial\Omega_2$, periodic boundary conditions are applied. The extent of the computational domain is chosen to be large enough in the wall-parallel directions that the convenient (though non-physical) periodic boundary conditions on $\partial\Omega_2$ have no effect on the nature of the near-wall turbulence, as illustrated qualitatively in figure 2.

The governing equations may be written functionally as

$$\mathcal{N}(U) = 0 \qquad \text{in } \Omega \tag{2.1a}$$

with boundary conditions

$$u_i = \Phi\, n_i \qquad \text{on } \partial\Omega_1 \tag{2.1b}$$

and prescribed initial conditions

$$u_i = u_i(0) \qquad \text{at } t = 0. \tag{2.1c}$$

For clarity, all differential equations will be written in operator form in this discussion, with these operators defined when first introduced. The (nonlinear) Navier-Stokes operator for the present case, in which the flow is assumed to have uniform density and viscosity, is given by

$$\mathcal{N}(U) = \begin{pmatrix} \dfrac{\partial u_i}{\partial t} + u_j \dfrac{\partial u_i}{\partial x_j} - \nu \dfrac{\partial^2 u_i}{\partial x_j^2} + \dfrac{1}{\rho}\dfrac{\partial p}{\partial x_i} + \dfrac{1}{\rho}\delta_{1i} P_x \\ \dfrac{\partial u_j}{\partial x_j} \end{pmatrix}.$$

In this discussion, x_1 is the streamwise direction, x_2 is the wall-normal direction, x_3 is the spanwise direction, the u_i's are the corresponding velocities, p is the pressure, ρ is the density, μ is the absolute viscocity, and $\nu \equiv \mu/\rho$ is the kinematic viscocity. The flow is sustained by a mean pressure gradient P_x in the streamise direction, which is modified at each time step in order to maintain a constant mass flux through the channel. Define also n as a wall-normal unit vector directed *into* the channel, $\overline{\tau_w} \equiv \mu\, \partial\overline{u_1}/\partial n|_{wall}$ as the mean skin friction on the wall for the uncontrolled channel (averaged in space and time), $u_\tau \equiv (\overline{\tau_w}/\rho)^{1/2}$ as the mean friction velocity, δ as the channel half-width, and $Re_\tau \equiv u_\tau \delta/\nu$ as the Reynolds number based on the mean friction velocity and the channel half width. The flow considered in this work is taken at $Re_\tau = 180$. All velocities are normalized by the friction velocity u_τ, and therefore may also be marked with a $(^+)$ superscript. All lengths are normalized by δ unless marked with a $(^+)$ superscript, in which case they are normalized by the wall unit ν/u_τ. All times are normalized by δ/u_τ unless marked with a $(^+)$ superscript, in which case they are normalized by ν/u_τ^2. Note that, with this normalization, $\nu = 1/Re_\tau$ in the above equation for $\mathcal{N}(U)$.

Three state vectors are used in this work: the flow U, the flow perturbation U', and the adjoint U^*:

$$U = \begin{pmatrix} u_i(x_1, x_2, x_3, t) \\ p\,(x_1, x_2, x_3, t) \end{pmatrix}, \quad U' = \begin{pmatrix} u_i'(x_1, x_2, x_3, t) \\ p'\,(x_1, x_2, x_3, t) \end{pmatrix}, \quad U^* = \begin{pmatrix} u_i^*(x_1, x_2, x_3, t) \\ p^*\,(x_1, x_2, x_3, t) \end{pmatrix}.$$

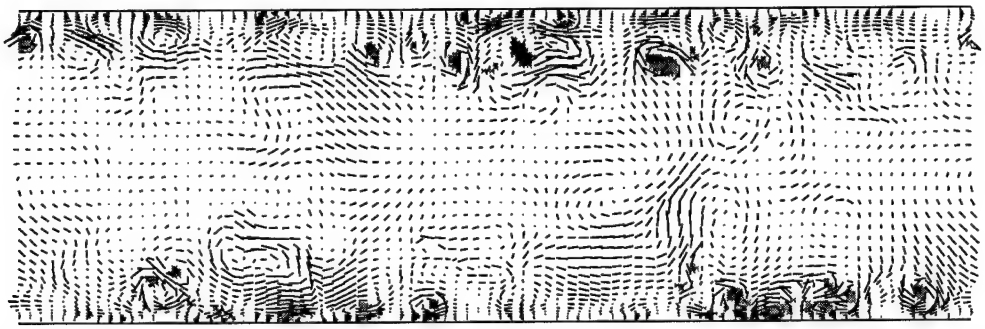
Note that each of these vector fields is composed of three velocity components and a pressure component. The motivation for introducing U' and U^* will be apparent in the derivation of the control equations to follow, and the parital differential equations governing these fields will be derived. Only after the control problem has been derived in differential form is it discretized in space and time. For the current three-dimensional nonlinear problem, this approach is found to yield systems of equations which are easiest to understand and to code. Note that for simpler systems of equations, such as the one-dimensional linear problem of transition control examined in Bewley, Agarwal, & Liu (1996), the matrix control equations derived from the discrete form of the governing equations is found to be tractable.

3. Analysis of control problem

The base turbulent flow analyzed in this problem is illustrated in figure 2. This flow has been carefully studied by Kim, Moin, & Moser (1987); the present simulations yield essentially the same second order statistics as the results presented there for the uncontrolled flow, though the numerical method used in the present work is substantially different in order to better facilitate wall-normal velocity boundary conditions.

As direct numerical simulations of turbulence produce a tremendous amount of data, it is important to analyze relevant statistics of these flowfields in order to better understand the phenomena taking place in an integrated sense and how these integral measures of the turbulence are modified by the addition of control. The statistics used to examine the turbulent flowfields in the present work are the mean velocity $\overline{u_1}$, the root-mean-square velocity fluctuations $u_{i,rms}$, the Reynolds stress $-\overline{u_1 u_2}$, the total stress $-\overline{u_1 u_2} + \nu\, \partial\overline{u_1}/\partial x_2$, and the two-point correlations $R_{ij}(\mathbf{r}) \equiv \overline{u_i(\mathbf{x})\, u_j(\mathbf{x}+\mathbf{r})}$ and their Fourier transform, the cospectra $E_{ij}(\mathbf{k})$. Note that the overbar $(\overline{\ })$ implies averages in the homogeneous directions x_1 and x_3 and, when appropriate, time. The statistics are functions of the nonhomogeneous direction x_2. A further discussion of these statistics and their behaviour in a turbulent channel at $Re_\tau = 180$ may be found in Kim, Moin, & Moser (1987).

(a) Three-dimensional visualization. For clarity, discriminant is marked only in the lower half of the domain. Flow is from left to right, walls are shaded.



(b) Crossflow visualization. For clarity, crossflow velocity vectors (indicated by the arrows) are marked on only one ninth of the gridpoints used in the computation.

FIGURE 2. Turbulent channel flow realization at $Re_\tau = 180$, no control. Regions of the flow with positive discriminant $D > D_{\text{threshold}}$ are shaded, indicating fluid motion which, in a pointwise sense, is vortical in nature, as suggested by Blackburn, Mansour, & Cantwell (1996). Small amounts of blowing and suction will be applied through the computational equivalent of closely spaced holes drilled in the walls in response to these turbulent motions in a manner which reduces drag.

The nature of the turbulent motion is also well characterized by observing the fluid at various points throughout the channel in a reference frame which moves with the local velocity. In this reference frame, the point under consideration is a critical point, as the local streamline slope is indeterminate. Thus, a critical point analysis of the type discussed by Perry & Chong (1987) is appropriate. Chong, Perry, & Cantwell (1990) and Blackburn, Mansour, & Cantwell (1996) have demonstrated that a single scalar quantity D, the discriminant of the velocity gradient tensor, provides a useful identification of regions in the flow which, in this context, are "focus" in nature. Such focus regions roughly correspond to "vortex-type" regions in a turbulent flowfield, though this description is only pointwise in nature.†

The velocity gradient tensor discussed in this work is defined in wall units $A_{ij} \equiv \partial u_i^+ / \partial x_j^+$. The second and third invariants of A are $Q \equiv \left\{[\text{tr}(A)]^2 - \text{tr}(A^2)\right\}/2$ and $R \equiv -\det(A)$. The discriminant of the velocity gradient tensor is given by $D = (27/4)R^2 + Q^3$. Regions with $D > 0$ are characterized by a velocity gradient tensor with one real and two complex eigenvalues (and thus a swirling, vortex-type motion in a Lagrangian reference frame), whereas regions with $D \leqslant 0$ are characterized by three real eigenvalues. For clarity, the visualizations of the discriminant presented in this work identify only regions

† Note that this description of a "vortex" is by no means unique. Robinson (1991) and Bernard, Thomas, & Handler (1993) discuss other vortex identification techniques.

of positive discriminant greater than a small threshold value $D > D_{\text{threshold}}$, where $D_{\text{threshold}} = 10^{-5}$.

4. Summary of optimal control theory

4.1. *Cost functional*

The first step in solving an optimal control problem is to represent the control problem of interest as a cost functional, $\mathcal{J}$, to be minimized. In the present problem, control is to be applied to minimize the drag averaged over a section of wall with area A and over the time period $(0, T]$ using the least amount of control effort possible. A relevant cost functional for the present problem is thus

$$\mathcal{J}(\Phi) = \frac{1}{AT} \int_w \int_0^T \left(\frac{\ell^2}{2} \Phi^2 + \mu \frac{\partial u_1}{\partial n} \right) dt\, dS.$$

The first term in the integrand is a measure of the magnitude of the control. The second term is a measure of exactly that quantity we would like to regulate: in this case, the drag. These quantities are integrated over the wall section under consideration, of area A, and over the time period under consideration, of duration T. Finally, they are weighted together with a factor $\ell^2/2$, which represents the price of the control. This quantity is small if the control is "cheap" (which reduces the significance of the first term), and large if applying control is "expensive".

4.2. *Gradient of cost functional*

As suggested by Abergel and Temam (1990), a rigorous procedure may be developed to determine the sensitivity of a cost functional $\mathcal{J}$ to small modifications of the control Φ for nonlinear problems of this sort. To do this, consider the perturbation to the cost functional resulting from a small perturbation to the control Φ in the direction Φ'. (Note that this control perturbation direction Φ' is arbitrary and scaled to have unit norm.) Define $\mathcal{J}'$ as the Fréchet differential (Vainberg 1964) of the cost functional such that

$$\mathcal{J}' \equiv \lim_{\epsilon \to 0} \frac{\mathcal{J}(\Phi + \epsilon\Phi') - \mathcal{J}(\Phi)}{\epsilon} \equiv \int_w \int_0^T \frac{\mathscr{D}\mathcal{J}(\Phi)}{\mathscr{D}\Phi} \Phi'\, dt\, dS.$$

The quantity $\mathcal{J}'$ is the cost functional perturbation due to a control perturbation $\epsilon\Phi'$ scaled by the inverse of the control perturbation magnitude ϵ in the limit that $\epsilon \to 0$. The above relation, considered for arbitrary Φ', also defines the gradient of the cost functional $\mathcal{J}$ with respect to the control Φ, which is written $\mathscr{D}\mathcal{J}(\Phi)/\mathscr{D}\Phi$.

In the present approach, the cost functional perturbation $\mathcal{J}'$ defined above is expressed as a simple linear function of the direction of the control perturbation Φ' through the solution of an adjoint problem. By the above formula, such a representation reveals the gradient direction $\mathscr{D}\mathcal{J}(\Phi)/\mathscr{D}\Phi$ directly. With this gradient information, the control Φ is updated on $(0, T]$ in the direction that, at least locally (*i.e.* for infinitesimal control updates), most effectively reduces the cost functional. The finite distance the control is updated in this direction is then found by a line search routine, which makes this iteration procedure very robust even when controlling nonlinear phenomena. The flow resulting from the modified control is then computed according to the (nonlinear) Navier-Stokes equation (1.1), the sensitivity of this new flow to further control modification is computed, and the process repeated. Upon convergence of this iteration, the flow is

advanced over the interval $(0, T_1]$, where $T_1 \leqslant T$, and an iteration for the optimal control over a new time interval $(T_1, T_1 + T]$ begins anew.†

The cost functional perturbation $\mathcal{J}'$ resulting from a control perturbation in the direction Φ' is given by

$$\mathcal{J}'(\Phi) = \int_w \int_0^T \frac{\mathscr{D}\mathcal{J}(\Phi)}{\mathscr{D}\Phi}\,\Phi'\,dt\,dS = \frac{1}{AT}\int_w \int_0^T \left(\ell^2\,\Phi\,\Phi' + \mu\,\frac{\partial u_1'}{\partial n}\right) dt\,dS, \tag{4.1}$$

where u_1' is the Fréchet differential of u_1, as defined in the following subsection. Adjoint calculus is used simply to re-express the integral involving u_1' as a linear function of Φ'. Once this is accomplished, Φ' is factored out of the integrands and, as the equation holds for arbitrary Φ', an expression for the gradient $\mathscr{D}\mathcal{J}(\Phi)/\mathscr{D}\Phi$ is extracted.

4.2.1. *Perturbation field*

Define U' as the Fréchet differential of U such that

$$U' \equiv \lim_{\epsilon \to 0} \frac{U(\Phi + \epsilon\Phi') - U(\Phi)}{\epsilon},$$

The quantity U' is the flow perturbation due to a control perturbation $\epsilon\Phi'$ scaled by the inverse of the control perturbation magnitude ϵ in the limit that $\epsilon \to 0$. The equations governing the dependence of the flow perturbation U' on the direction of the control perturbation Φ' may be found by taking the Fréchet differential of the state equation (1.1) itself. The result is

$$\mathcal{N}'\,U' = 0, \qquad \text{in } \Omega \tag{4.2a}$$

with boundary conditions

$$u_i' = \Phi'\,n_i \qquad \text{on } \partial\Omega_1 \tag{4.2b}$$

and initial conditions

$$u_i' = 0 \qquad \text{at } t = 0, \tag{4.2c}$$

where the differential Navier-Stokes operator $\mathcal{N}'$ is given by

$$\mathcal{N}'\,U' = \begin{pmatrix} \dfrac{\partial u_i'}{\partial t} + u_j\dfrac{\partial u_i'}{\partial x_j} + u_j'\dfrac{\partial u_i}{\partial x_j} - \nu\dfrac{\partial^2 u_i'}{\partial x_j^2} + \dfrac{1}{\rho}\dfrac{\partial p'}{\partial x_i} \\ -\dfrac{1}{\rho}\dfrac{\partial u_j'}{\partial x_j} \end{pmatrix}. \tag{4.3}$$

Note that the operation $\mathcal{N}'\,U'$ is linear in the perturbation field U', though the operator $\mathcal{N}'$ itself is a function of the solution U of the Navier-Stokes problem. Equation (4.2) reflects the linear dependence of the perturbation field U' in the interior of the domain

† Note that the flow may be advanced over a time interval which is shorter than the complete interval over which the optimization was performed. The rationale for such an approach is that the controls computed near the end of each optimization interval are computed without regard to the (inevitable) further development of the flow beyond the end of the optimization interval. Thus, the controls near the end of the optimization interval are not as effective in the long run as those controls near the beginning of the interval, which are optimized with greater "foresight" about the flow development. Thus, the controls optimized near the end of one interval may be improved upon (in terms of the long-term performance) by recalculation in the following interval.

on the direction of the control perturbation Φ' at the boundary. However, the implicit linear relationship $U' = U'(\Phi')$ given by these equations is not tractable for expressing $\mathcal{J}'$ in a form from which $\mathscr{D}\mathcal{J}(\Phi)/\mathscr{D}\Phi$ may be deduced. For the purpose of determining a more useful relationship with which we may express $\mathcal{J}'$ in the desired form, we now appeal to an adjoint identity.

4.2.2. *Derivation of adjoint identity*

This subsection derives the adjoint of the linear partial differential operator $\mathcal{N}'$.

Define an inner product over the domain in space-time under consideration such that

$$\langle U', U^* \rangle = \int_\Omega \int_0^T U' \cdot U^* \, dt \, dV,$$

and consider the identity

$$\boxed{\langle \mathcal{N}' U', U^* \rangle = \langle U', \mathcal{N}^* U^* \rangle + b.} \tag{4.4}$$

Integration by parts may be used to move all differential operations from U' on the left hand side of the equation to U^* on the right hand side, resulting in the definition of the adjoint operator

$$\mathcal{N}^* U^* = \begin{pmatrix} -\dfrac{\partial u_i^*}{\partial t} - u_j \Big(\dfrac{\partial u_i^*}{\partial x_j} + \dfrac{\partial u_j^*}{\partial x_i} \Big) - \nu \dfrac{\partial^2 u_i^*}{\partial x_j^2} + \dfrac{1}{\rho} \dfrac{\partial p^*}{\partial x_i} \\ -\dfrac{1}{\rho} \dfrac{\partial u_j^*}{\partial x_j} \end{pmatrix}, \tag{4.5}$$

where the operator $\mathcal{N}^*$ is a function of U, and an expression for b, which contains all the boundary terms:

$$b = \int_w \int_0^T -n_j \left(u_j \, u_i' \, u_i^* + u_i \, u_j' \, u_i^* - \nu \Big(\frac{\partial u_i'}{\partial x_j} \, u_i^* - u_i' \frac{\partial u_i^*}{\partial x_j} \Big) + \frac{1}{\rho} \Big(p' \, u_j^* - u_j' \, p^* \Big) \right) dt \, dS$$
$$+ \int_\Omega u_i' \, u_i^* \, dV \Big|_{t=T} - \int_\Omega u_i' \, u_i^* \, dV \Big|_{t=0}.$$

The identity (4.4) is very powerful; in fact, simplification of this identity by interior equations, boundary conditions, and initial conditions on U, U', and U^* provides an expression which recasts $\mathcal{J}'$ in a form from which $\mathscr{D}\mathcal{J}(\Phi)/\mathscr{D}\Phi$ may be deduced, as illustrated in the following two subsections.

4.2.3. *Definition of adjoint field*

Consider an adjoint state defined (as yet, arbitrarily) by

$$\mathcal{N}^* U^* = 0, \qquad \text{in } \Omega \tag{4.6a}$$

with boundary conditions

$$u_i^* = \delta_{i1} \qquad \text{on } \partial\Omega_1 \tag{4.6b}$$

and initial conditions

$$u_i^* = 0 \qquad \text{at } t = T, \tag{4.6c}$$

where the adjoint operator $\mathcal{N}^*$ is given in (4.5). Note that the adjoint problem (1.2), though linear, has complexity similar to that of the Navier-Stokes problem (1.1). Note

also that the "initial" conditions in (4.6c) are defined at $t = T$. With this definition, the adjoint field must be marched *backward* in time over the interval—due to the sign of the time derivative and viscous terms in the adjoint operator $\mathcal{N}^*$ in (4.5), this is the natural direction for this time march. However, as the operator $\mathcal{N}^*$ is a function of U, computation of the adjoint field U^* requires storage of the flow field U on $t \in [0, T]$, which itself must be computed with a *forward* march. This storage issue presents a numerical complication which precludes solution for large optimization invervals T. However, the problem is not insurmountable for moderate values of T.

Equation (1.2) is, as yet, simply a *definition* of an adjoint field. The motivation for considering an adjoint field so defined is revealed in the following subsection.

4.2.4. *Identification of gradient*

The identity (4.4) is now simplified using the equations defining the state field (1.1), the perturbation field (4.2), and the adjoint field (1.2). Due to the judicious choice of RHS forcing terms in the equations (1.2a)–(4.6c) defining the adjoint field, the identity reduces to

$$\int_w \int_0^T \mu \frac{\partial u_1'}{\partial n} \, dt \, dS = - \int_w \int_0^T p^* \, \Phi' \, dt \, dS.$$

Using this equation, the cost functional perturbation $\mathcal{J}'$ in (4.1) may be conveniently rewritten as

$$\int_w \int_0^T \left(\frac{\mathscr{D}\mathcal{J}(\Phi)}{\mathscr{D}\Phi} - \ell^2 \, \Phi + p^* \right) \Phi' \, dt \, dS = 0.$$

As Φ' is arbitrary, this implies that

$$\frac{\mathscr{D}\mathcal{J}(\Phi)}{\mathscr{D}\Phi} = \ell^2 \, \Phi - p^*. \tag{4.7}$$

Thus, the desired gradient is found to be a function of the solution of the adjoint problem (1.2) discussed in §4.2.3.

4.3. *Gradient update to control*

4.3.1. *Simple gradient*

A control strategy using a simple gradient (also known as steepest descent) algorithm may now be proposed such that

$$\Phi^k = \Phi^{k-1} - \alpha \frac{\mathscr{D}\mathcal{J}(\Phi^{k-1})}{\mathscr{D}\Phi}, \tag{4.8}$$

over the entire time interval $t \in (0, T]$, where k indicates the iteration number and α is a parameter of descent which governs how large an update is made, which is adjusted at each iteration step to be that value which minimizes $\mathcal{J}$. This algorithm updates Φ at each iteration in the direction of maximum decrease of $\mathcal{J}$. As $k \to \infty$, the algorithm should converge to some local minimum of $\mathcal{J}$ over the domain of the control Φ on the time interval $t \in (0, T]$. Note that convergence to a global minimum will not in general be attained by such a scheme, and that, as time proceeds, $\mathcal{J}$ will not necessarily decrease.

4.3.2. *Conjugate gradient*

The simple gradient approach described above is straightforward, but, as illustrated in figure 3, not always efficient. Even in linear problems, for cases in which the cost functional has a long, narrow "valley", the lack of a momentum term from iteration to

iteration tends to cause the simple gradient algorithm to bounce from one side of the valley to the other without turning to proceed along the valley floor. The conjugate gradient approach tends to improve this behaviour, and will be examined further in future work.

5. Numerical approach

5.1. *Numerical algorithm for solution of flow and adjoint equations*

The control formulations derived above were tested in direct numerical simulations of fully developed turbulent channel flows at $Re_\tau = 180$. Fourier transforms are used to compute spatial derivatives in the homogeneous directions with 3/2 dealiasing on the nonlinear terms, and a conservative second-order finite difference scheme is used to compute spatial derivatives in the wall-normal direction. For the present simulations, the number of Fourier modes used is $170 \times 129 \times 170$ in the x_1, x_2, and x_3 directions respectively (*i.e.* $256 \times 129 \times 256$ dealiased collocation points), and the size of the computational domain in wall units is $L_1^+ = 2260$, $L_2^+ = 360$, $L_3^+ = 1130$. The resulting effective grid resolution in the streamwise and spanwise directions (on collocation points determined without the extra 3/2 padding) is $\Delta x_1^+ = 13$, $\Delta x_3^+ = 7$. Hyperbolic tangent streching of the grid is used in the wall-normal direction, resulting in a grid spacing of $\Delta x_2^+ = 0.3$ adjacent to the wall and $\Delta x_2^+ = 5$ in the center of the channel. Fine grid resolution is required near the wall to resolve the shear layer; the mesh is fine in this direction even up to the center of the domain because the second order difference scheme used to compute the derivatives in this direction is numerically dissipative. The computational grid is staggered in the wall-normal direction to prevent decoupling of the even and odd modes of the pressure.

The flow is advanced in time using an explicit low-storage third-order Runge-Kutta method for terms involving x_1 and x_3 derivatives and an implicit Crank-Nicholson method at each Runge-Kutta substep for x_2 derivative terms. A temporal discretization implicit in the x_2 derivatives is necessary to mitigate the CFL time step restriction when control is applied, as the control fluid at the wall is directed in the x_2 direction, which is precisely the region and direction in which the mesh must be refined most to resolve the shear layer.

The adjoint solver is coded with a method analogous to that of the flow solver. The flow field is stored every 5 time steps on the forward sweep, with linear interpolation of these stored fields used on the backward sweep to determine the operator $\mathcal{N}^*$. In the optimal calculations presented here, we chose $\ell = 10^{-2}$ (control power is taken to be "cheap"). The Polak-Ribiere variant of the conjugate gradient algorithm was used for the control update, with α computed at each iteration by Brent's method, a robust line minimization algorithm taken from Press *et al.* (1986).

6. Performance of controlled systems

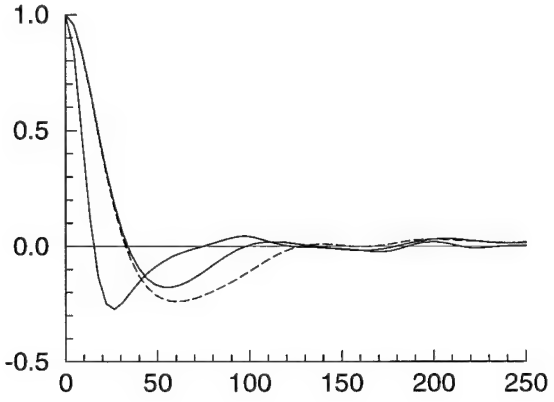
At the time this document went to press, high Reynolds number computational results were not yet available. They will appear in the version of this paper that gets submitted to *J. Fluid Mech.* Instead, shown here are some of the visualizations and statistics from the baseline case (no control) which was run in order validate the code and illustrate the post processing analysis to be performed.

Figure 6 demonstrates that effective performance, including about 50% drag reduction and nearly an order of magnitude reduction in the turbulent kinetic energy, have been obtained in $Re_\tau = 100$ turbulent flow.

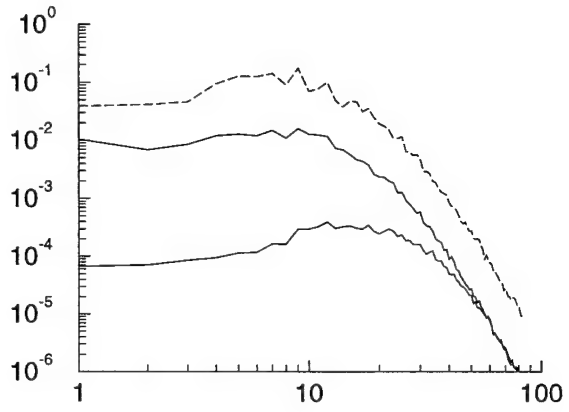
(a) Top view. For clarity, discriminant is marked only in lower half of the domain.

(a) Side view. For clarity, $D_{threshold}$ has been increased to 5×10^{-5}.

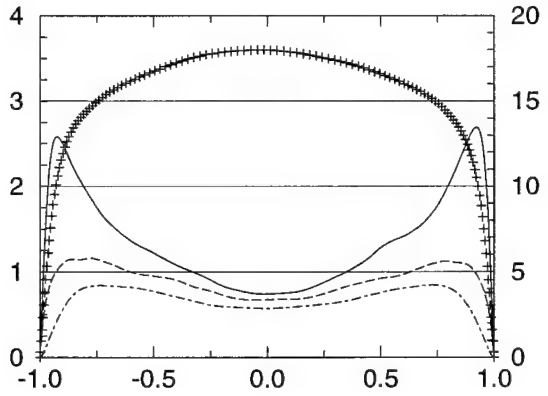
FIGURE 3. Visualization of discriminant in unmanipulated turbulent channelf flow. Areas of positive discriminant $D > D_{\text{threshold}}$ are shaded, indicating fluid motion which, in a pointwise sense, is vortical in nature.



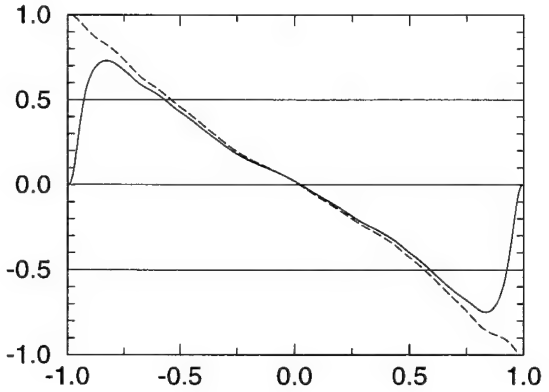
(a) Spanwise two-point correlations at $x_2^+ = 5.4$*:* ——, R_{11}; - - - -, R_{22}; — — —, R_{33}.



(b) Spanwise one-dimensional energy spectra at $x_2^+ = 5.4$*:* ——, E_{11}; - - - -, E_{22}; — — —, E_{33}.

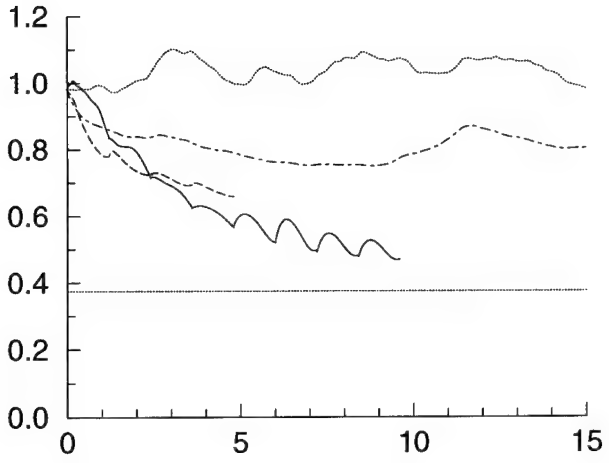


(c) Profiles of RMS velocity fluctuations (scale at left): ——, $u_{1,rms}$; - - - -, $u_{3,rms}$; — — —, $u_{2,rms}$, *and mean velocity (scale at right):* +, $\overline{u_1}$.

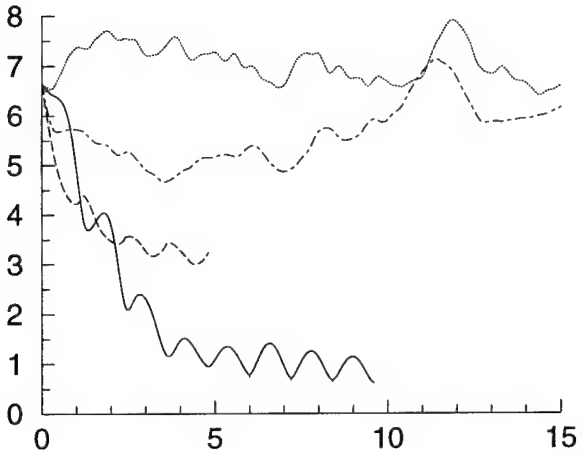


(d) Profiles of Reynolds stress and total stress: ——, *Reynolds stress* $-\overline{u_1 u_2}$; - - - -, *total stress* $-\overline{u_1 u_2} + \nu\, \partial \overline{u_1}/\partial x_2$.

FIGURE 4. Selected instantaneous statistics of the unmanipulated turbulent flow.



(a) History of drag.



(b) History of turbulent kinetic energy.

FIGURE 5. Performance of optimal control algorithm at $Re_\tau = 100$: ---- , optimal drag formulation (§4); ——— , optimal energy formulation; — — — , intuition based scheme $\Phi = -u_2(y^+ = 10)$; ····· , no control (both turbulent and laminar drag curves are shown).

REFERENCES

ABERGEL, F. & TEMAM, R. 1990 On some control problems in fluid mechanics. *Theor. and Comp. Fluid Dynamics* **1**, 303-325.

AUBRY, N., HOLMES, P., LUMLEY, J.L., & STONE, E. 1988 The dynamics of coherent structures in the wall region of a turulent boundary layer. *J. Fluid Mech.* **192**, 115-173.

BEWLEY, T.R., AGARWAL, R., & LIU, S. 1996 Optimal and robust control of transition. Submitted to *J. Fluid Mech.*

BERKOOZ, G., HOLMES, P., & LUMLEY, J.L. 1993 The proper orthogonal decomposition in the analysis of turbulent flows. *Ann. Rev. Fluid Mech.* **25**, 539-575.

BERNARD, P.S., THOMAS, J.M., & HANDLER, R.A. 1993 Vortex dynamics and the production of Reynolds stress. *J. Fluid Mech.* **253**, 385-419.

BEWLEY, T.R., MOIN, P., & TEMAM, R. 1996 A method for optimizing feedback control rules for wall-bounded turbulent flows based on control theory. *Proceedings of the ASME Fluids*

Engineering Summer Meeting, ASME FED **237**, p. 279-285.

BLACKBURN, H.M., MANSOUR, N.N., & CANTWELL, B.J. 1996 Topology of fine-scale motions in turbulent channel flow. *J. Fluid Mech.* **310**, 269-292.

CHOI, H., MOIN, P., & KIM, J. 1994 Active turbulence control for drag reduction in wall-bounded flows. *J. Fluid Mech.* **262**, 75-110.

CHONG, M.S., PERRY, A.E., & CANTWELL B.J. 1990 A general classification of three-dimensional flow fields. *Phys Fluids* A **2**, 765-777.

COLLER, B.D., HOLMES, P., & LUMLEY, J.L. 1994 Control of bursting in boundary layer models. *Appl. Mech. Rev.* **47** (6), part 2, S139-S143.

COLLER, B.D., HOLMES, P., & LUMLEY, J.L. 1994 Control of noisy heteroclinic cycles. *Physica* D **72**, 135-160.

DOYLE, J.C., GLOVER, K., KHARGONEKAR, P.P., & FRANCIS, B.A. 1989 State-space solutions to standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems. *IEEE Trans. Auto. Control* **34**, 8, 831-847.

FINLAYSON, B.A. 1972 *The Method of Weighted Residuals and Variational Principles.* Academic.

GUNZBURGER, M.D., HOU, L., & SVOBODNY, T.P. 1990 A numerical method for drag minimization via the suction and injection of mass through the boundary. In *Stabilization of Flexible Structures* (ed. J.P. Zolesio). Springer.

HERTZ, J., KROGH, A., & PALMER, R.G. 1991 *Introduction to the Theory of Neural Computation.* Addison-Wesley.

KEEFE, L.R., MOIN, P., & KIM, J. 1992 The dimension of attractors underlying periodic turbulent Poiseuille flow. *J. Fluid Mech.* **242**, 1-29.

KIM, J., MOIN, P., & MOSER, R. 1987 Turbulence statistics in fully developed channel flow at low Reynolds number. *J. Fluid Mech.* **177**, 133-166.

LEE, C., KIM, J., BABCOCK, D., & GOODMAN, R. Application of neural network to turbulence control for drag reduction. Submitted to *Phys. Fluids.*

LIONS, J.L. 1968 *Controle Optimal des Systèmes Gouvernés par des Equations aux Dérivées Partielles.* Dunod. (English translation, Springer.)

LUENBERGER, D.G. 1969 *Optimization by Vector Space Methods.* Wiley.

MCMICHAEL, J.M. 1996 Progress and prospects for active flow control using Microfabricated Electro-Mechanical Systems (MEMS). *AIAA Paper No.* 96-0306.

MOIN, P, & BEWLEY, T 1994 Feedback control of turbulence. *Appl. Mech. Rev.* **47** (6), part 2, S3-S13.

MORSE, P.M. & FESHBACH, H. 1953 *Methods of Theoretical Physics.* McGraw-Hill.

PERRY, A.E. & CHONG, M.S. 1987 A description of eddying motions and flow patterns using critical-point concepts. *Ann. Rev. Fluid Mech.* **19**, 125-55.

PRESS, W.H., FLANNERY, B.P., TEUKOLSKY, S.A., & VETTERLING, W.T. 1986 *Numerical Recipes.* Cambridge.

ROBINSON, S.K. 1991 Coherent motions in the turbulent boundary layer. *Ann. Rev. Fluid Mech.* **23**, 601-639.

SRITHARAN, S. 1991 Dynamic programming of the Navier-Stokes equations. *Systems and Control Letters* **16**, (4), 299-307.

VAINBERG, M. 1964 *Variational Methods for the Study of Nonlinear Operators.* Holden-Day.

Intentionally left blank

PART C.

Robust control of turbulence

The framework for applying optimal and robust control theories to linear problems is first reviewed in a notation fairly consistent with that used by Doyle *et al.* (1989). The discussion is made strictly in the time domain, not the frequency domain often used to discuss these approaches, to facilitate extension of the approaches to nonlinear problems in which frequency domain techniques are of limited usefulness. The resulting development is fairly straightforward and does not assume the reader is accustomed to the language of control theory.

Nonlinear optimal control problems are then reviewed in a similar notation—the approach described here is that taken by Bewley, Moin, and Temam (1997) for the optimal control of wall-bounded turbulent flows. Finally, the concepts of robust control are extended to nonlinear problems, such as the control of turbulence, in a consistent manner.

---

1. Outline of linear regulation problems

1.1. *Optimal regulation of linear problems*

1.1.1. *State equation*

Consider a state vector $\mathbf{u}$ which is a function of some feedback control vector Φ such that it obeys the linear evolution equation

$$\boxed{\dot{\mathbf{u}} = A\,\mathbf{u} + B_2\,\Phi} \tag{1.1a}$$

with given initial conditions

$$\mathbf{u} = \mathbf{u}(0) \qquad \text{at } t = 0. \tag{1.1b}$$

The matrices A and B_2 may be functions of time but do not themselves depend on the state $\mathbf{u}$ or the control Φ.

1.1.2. *Cost function*

The object of applying control in the present problem is to regulate some measure of the state to zero quickly without applying excessive amounts of control. Mathematically, this objective is expressed as the minimization of a cost functional which balances a measure of the state $\mathbf{u}$ with a measure of the control Φ applied. We will use the norm symbol to denote these measures, which may be defined appropriately for particular problems of interest:

$$\mathcal{J}_2 \equiv \frac{1}{2\,T} \int_0^T \left(||\mathbf{u}||^2 + \ell^2 \, ||\Phi||^2 \right) \, dt.$$

Note that the two terms are weighted together with a factor ℓ^2 which accounts for the "price" of the control; this factor is large if applying the control is "expensive", which emphasizes the importance of the latter term in this equation and generally results in a modest control effort, and small if applying the control is "cheap", which results in a

larger control effort. These terms are then averaged over some optimization interval of consideration $[0,T]$. In matrix form, $\mathcal{J}_2$ is expressed as

$$\boxed{\mathcal{J}_2 = \frac{1}{2T}\int_0^T \left(\mathbf{u}^* C_1^* C_1 \mathbf{u} + \ell^2 \Phi^* \Phi\right) dt}$$

with C_1 defined appropriately based on the definition of the norm of the state $||\mathbf{u}||$ and the star (*) denoting the conjugate transpose. By appropriate scaling of the vector Φ and the matrix B_2, the norm of the control $||\Phi||$ is taken simply as the Euclidian norm without loss of generality.

A technique to design a feedback control relationship of the form $\Phi = K_2 \mathbf{u}$ which minimizes the cost function $\mathcal{J}_2$ is now briefly outlined.

1.1.3. *Adjoint equation*

Define an adjoint state (as yet, arbitrarily) by the relation

$$-\dot{\lambda} = A^* \lambda + C_1^* C_1 \mathbf{u} \tag{1.2a}$$

with initial conditions

$$\lambda = 0 \qquad \text{at } t = T. \tag{1.2b}$$

Note that the "initial" conditions (1.2b) are defined at $t = T$, so to determine the adjoint on the interval $[0,T)$, the evolution equation (1.2a) must be marched *backwards* from $T \to 0$.

1.1.4. *Gradient of cost function $\mathcal{J}_2$ with respect to control Φ*

It may be shown that the gradient of the cost with respect to the control is a simple function of the adjoint state defined by (1.2):

$$\frac{\mathscr{D}\mathcal{J}_2}{\mathscr{D}\Phi} = \ell^2 \Phi + B_2^* \lambda. \tag{1.3}$$

1.1.5. *Solution of control problem*

By (1.3), the most suitable control which results in

$$\frac{\mathscr{D}\mathcal{J}_2}{\mathscr{D}\Phi} = 0 \quad \text{(minimum)} \tag{1.4}$$

as a function of the adjoint state is given simply by

$$\Phi = -\frac{1}{\ell^2} B_2^* \lambda \tag{1.5}$$

Combining the state equation (1.1a), the adjoint equation (1.2a), and the control given by (1.5) into a combined matrix form gives

$$\begin{bmatrix} \dot{\mathbf{u}} \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} A & -\frac{1}{\ell^2} B_2 B_2^* \\ -C_1^* C_1 & -A^* \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \lambda \end{bmatrix} \tag{1.6}$$

Now prescribe a relationship between any state vector $\mathbf{u} = \mathbf{u}(t)$ and the corresponding adjoint $\lambda = \lambda(t)$ such that

$$\lambda = X_2 \mathbf{u}, \tag{1.7}$$

where $X_2 = X_2(t)$. Inserting this expression into (1.6) to eliminate λ and combining the top and bottom rows to eliminate $\dot{\mathbf{u}}$ leads to the expression

$$\left(-\dot{X}_2 = A^* X_2 + X_2 A - X_2 \frac{1}{\ell^2} B_2 B_2^* X_2 + C_1^* C_1 \right) \mathbf{u}$$

As this expression is valid for any state vector $\mathbf{u}$, we arrive at a Riccati equation for the matrix $X_2(t)$:

$$-\dot{X}_2 = A^* X_2 + X_2 A - X_2 \frac{1}{\ell^2} B_2 B_2^* X_2 + C_1^* C_1 \tag{1.8a}$$

with initial conditions, due to (1.2b) and (1.7), given by

$$X_2 = 0 \qquad \text{at } t = T. \tag{1.8b}$$

Combining (1.5) and (1.7), the optimal control Φ as a function of the state $\mathbf{u}$ is given by the state feedback relationship

$$\boxed{\Phi = K_2 \mathbf{u} \qquad \text{where} \qquad K_2 = -\frac{1}{\ell^2} B_2^* X_2} \tag{1.9}$$

where $X_2(t)$ is the solution of (1.8) and thus $K_2 = K_2(t)$.

1.1.6. *Infinite time horizon for time invariant problems*

If the matrices A, B_2, and C_1 are time invariant, then in the limit of large optimization intervals $T \to \infty$ the matrix $X_2(t)$ defined by (1.8) approaches a steady state value in the march from the initial conditions defined at $t = T$ back towards $t = 0$. This steady state value may be found by setting $\dot{X}_2 = 0$ in (1.8a), which leads to

$$\boxed{0 = A^* X_2 + X_2 A - X_2 \frac{1}{\ell^2} B_2 B_2^* X_2 + C_1^* C_1} \tag{1.10}$$

The optimum feedback relationship given by (1.9) in this limit is thus time invariant and a function of the solution to (1.10), referred to as an algebraic Riccati equation (ARE). Solution methods for equations of this type are well developed (Laub 1991).

1.2. *Simple interpretation of the adjoint field*

In the preceding discussion, the determination of optimal feedback control relationship $\Phi = K_2 \mathbf{u}$ in (1.9) was closely linked to the definition of an adjoint λ in (1.2). However, the definition of λ was made arbitrarily in (1.2), and subsequently justified only mathematically in (1.3) as being that field which is required to express the gradient of the cost function with respect to the control $\mathscr{D}\mathcal{J}_2/\mathscr{D}\Phi$ in a simple manner.

In the case that the control Φ enters the state equation (1.1) through the identity matrix, $B_2 = I$, a simple interpretation of the adjoint is now clear. In this case, the expression for the gradient (1.3) reduces to

$$\frac{\mathscr{D}\mathcal{J}_2}{\mathscr{D}\Phi} = \ell^2\,\Phi + \lambda. \tag{1.11}$$

Thus, the gradient consists of two terms. The first term simply accounts for the term in the cost function $\mathcal{J}_2$ which measures the magnitude of the control; in the absense of other term in the cost function, this term would drive the control to zero when $\mathcal{J}_2$ is minimized.

The second term in (1.11) accounts for the term in the cost function $\mathcal{J}_2$ which measures the state u itself. Thus, one interpretation of the adjoint λ is simply that: *The adjoint, when properly defined, is a measure of the sensitivity of the term of the cost function which measures the state u to additional RHS forcing of the state equation.* Note that there are exactly as many components of the adjoint λ as there are components of the state equation (1.1a).

1.3. *Robust regulation of linear problems*

1.3.1. *State equation*

Consider the linear state equation of (1.1) with additional forcing due to an external disturbance χ

$$\boxed{\dot{\mathbf{u}} = A\,\mathbf{u} + B_1\,\chi + B_2\,\Phi} \tag{1.12a}$$

with given initial conditions

$$\mathbf{u} = \mathbf{u}(0) \qquad\qquad \text{at } t = 0. \tag{1.12b}$$

The matrix B_1 may be a function of time but does not itself depend on the state $\mathbf{u}$ or the control Φ.

1.3.2. *Cost function*

The object of applying control in the robust problem is identical to the optimal problem, except we now play the "devil's advocate" and seek to find the best control in the presence of a "small" component of exactly that disturbance χ which is maximally aggrevating to the control objective. To represent this concept mathematically, we append to the cost function discussed in the previous section a term which accounts for the magnitude of the disturbance used to aggrevate the system

$$\mathcal{J}_\infty \equiv \frac{1}{2\,T}\int_0^T \left(||\mathbf{u}||^2 + \ell^2\,||\Phi||^2 - \gamma^2\,||\chi||^2\right)\,dt.$$

Note that the sign of the term which is used to account for the disturbance is opposite to the sign used to account for the control; this is because we *minimize* with respect the control Φ while simultaneously we *maximize* with respect to the disturbance χ. The term involving $-\gamma^2\,||\chi||^2$ limits the magnitude of the disturbance in the maximization with respect to χ as the term involving $\ell^2\,||\Phi||^2$ limits the magnitude of the control in the minimization with respect to Φ. In matrix form, $\mathcal{J}_\infty$ is expressed as

$$\boxed{\mathcal{J}_\infty = \frac{1}{2\,T}\int_0^T \left(\mathbf{u}^*\,C_1^*\,C_1\,\mathbf{u} + \ell^2\,\Phi^*\,\Phi - \gamma^2\chi^*\,\chi\right)\,dt}$$

By appropriate scaling of the vector χ and the matrix B_1, the norm of the disturbance $||\chi||$ is taken simply as the Euclidian norm without loss of generality.

A technique to design a feedback control relationship of the form $\Phi = K_\infty \mathbf{u}$ which minimizes the cost function $\mathcal{J}_\infty$ in the presence of a small component of the worst external disturbance χ forcing the state equation (1.12) is now briefly outlined. By designing a feedback control rule effective for a state disturbed in this manner, the control rule which is found is effective in the presence of small disturbances of any type, and has nearly the same nominal performance (*i.e.* performance on the undisturbed system) as the optimal controller determined in the previous section.

1.3.3. *Adjoint equation*

Define an adjoint state as for the optimal control case by the relation

$$-\dot{\lambda} = A^* \lambda + C_1^* C_1 \mathbf{u} \tag{1.13a}$$

with initial conditions

$$\lambda = 0 \qquad \text{at } t = T. \tag{1.13b}$$

1.3.4. *Gradients of cost function $\mathcal{J}_\infty$ with respect to control Φ and disturbance χ*

In a manner identical to the derivation leading to (1.3), the gradient of the cost with respect to the control Φ and the disturbance χ in this problem are simple functions of the adjoint state defined by (1.13):

$$\frac{\mathscr{D}\mathcal{J}_\infty}{\mathscr{D}\Phi} = \ell^2 \Phi + B_2^* \lambda \qquad \text{and} \qquad \frac{\mathscr{D}\mathcal{J}_\infty}{\mathscr{D}\chi} = -\gamma^2 \chi + B_1^* \lambda. \tag{1.14}$$

1.3.5. *Solution of control problem*

By (1.14), the most suitable control and disturbance which result in

$$\frac{\mathscr{D}\mathcal{J}_\infty}{\mathscr{D}\Phi} = 0 \ \text{(minimum)} \qquad \text{and} \qquad \frac{\mathscr{D}\mathcal{J}_\infty}{\mathscr{D}\chi} = 0 \ \text{(maximum)} \tag{1.15}$$

are given simply by

$$\Phi = -\frac{1}{\ell^2} B_2^* \lambda \qquad \text{and} \qquad \chi = \frac{1}{\gamma^2} B_1^* \lambda. \tag{1.16}$$

Combining the state equation (1.12a) and the adjoint equation (1.13a) with the control and disturbance given by (1.16) into a combined matrix form gives

$$\begin{bmatrix} \dot{\mathbf{u}} \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} A & \frac{1}{\gamma^2} B_1 B_1^* - \frac{1}{\ell^2} B_2 B_2^* \\ -C_1 C_1^* & -A^* \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \lambda \end{bmatrix} \tag{1.17}$$

Now prescribe a relationship between any state vector $\mathbf{u}$ and the corresponding adjoint λ such that

$$\lambda = X_\infty \mathbf{u} \tag{1.18}$$

Inserting this expression into (1.17) to eliminate λ and combining the top and bottom rows to eliminate $\dot{\mathbf{u}}$ leads to the expression

$$\left[-\dot{X}_\infty = A^* X_\infty + X_\infty A + X_\infty \left(\frac{1}{\gamma^2} B_1 B_1^* - \frac{1}{\ell^2} B_2 B_2^* \right) X_\infty + C_1^* C_1 \right] \mathbf{u}$$

As this expression is valid for any state vector $\mathbf{u}$, we arrive at a Riccati equation for the matrix $X_\infty(t)$:

$$-\dot{X}_\infty = A^* X_\infty + X_\infty A + X_\infty \left(\frac{1}{\gamma^2} B_1 B_1^* - \frac{1}{\ell^2} B_2 B_2^* \right) X_\infty + C_1^* C_1 \tag{1.19a}$$

with initial conditions, due to (1.13b) and (1.18), given by

$$X_\infty = 0 \qquad \text{at } t = T. \tag{1.19b}$$

Combining (1.16) and (1.18), a robust control Φ which is effective even in the presence of a small component of the worst case disturbance χ is given by the state feedback relationship

$$\boxed{\Phi = K_\infty \mathbf{u} \qquad \text{where} \qquad K_\infty = -\frac{1}{\ell^2} B_2^* X_\infty} \tag{1.20}$$

where X_∞ is the solution of (1.19) and thus $K_\infty = K_\infty(t)$.

1.3.6. *Infinite time horizon for time invariant problems*

If the matrices A, B_1, B_2, and C_1 are time invariant, then in the limit that the optimization interval $T \to \infty$ the matrix $X_\infty(t)$ defined by (1.19) approaches a steady state value in the march from the initial conditions defined at $t = T$ back towards $t = 0$, and is given by the solution to

$$\boxed{0 = A^* X_\infty + X_\infty A + X_\infty \left(\frac{1}{\gamma^2} B_1 B_1^* - \frac{1}{\ell^2} B_2 B_2^* \right) X_\infty + C_1^* C_1} \tag{1.21}$$

The robust feedback relationship given by (1.20) in this limit is thus time invariant and a function of the solution to (1.21).

2. Outline of nonlinear regulation problems

2.1. *Optimal regulation of nonlinear problems*

2.1.1. *State equation*

Consider a state vector $\mathbf{u}$ which is a function of some feedback control vector Φ such that it obeys the nonlinear evolution equation

$$\boxed{\dot{\mathbf{u}} = \mathcal{A}(\mathbf{u}) + \mathcal{B}_2(\Phi)} \tag{2.1a}$$

with given initial conditions

$$\mathbf{u} = \mathbf{u}(0) \qquad \text{at } t = 0. \tag{2.1b}$$

The nonlinear functions $\mathcal{A}(\mathbf{u})$ and $\mathcal{B}_2(\Phi)$ may themselves be functions of time.

2.1.2. *Cost function*

The object of applying control in the present case is identical to the optimal linear regulation problem described in §1.1.2. Mathematically, it is expressed as the minimization of

$$\boxed{\mathcal{J}_2 \equiv \frac{1}{2T} \int_0^T \left(\mathbf{u}^* C_1^* C_1 \mathbf{u} + \ell^2 \Phi^* \Phi \right) dt} \tag{2.2}$$

A technique to determine the control Φ on the interval $(0, T]$ which (locally) minimizes the cost function $\mathcal{J}_2$ for the nonlinear state equation (2.1) is now briefly outlined.

2.1.3. *Perturbation equation*

Consider the linear problem of a small perturbation $(\Phi', \mathbf{u}')$ to some reference solution $(\Phi, \mathbf{u})$ of the system given by (2.1). It is easily shown that such a perturbation must obey a linear evolution equation of the form

$$\boxed{\dot{\mathbf{u}}' = A\,\mathbf{u}' + B_2\,\Phi'} \tag{2.3a}$$

with initial conditions

$$\mathbf{u}' = 0 \qquad \text{at } t = 0. \tag{2.3b}$$

The matrices A and B_2 are functions of time and depend explicitly on the reference condition $(\Phi, \mathbf{u})$.

2.1.4. *Adjoint equation*

An adjoint system is defined based on the A matrix of the perturbation problem (2.3) such that

$$-\dot{\lambda} = A^*\,\lambda + C_1^*\,C_1\,\mathbf{u} \tag{2.4a}$$

with initial conditions

$$\lambda = 0 \qquad \text{at } t = T. \tag{2.4b}$$

2.1.5. *Gradient of cost function $\mathcal{J}_2$ with respect to control Φ*

As in the linear case, the gradient of the cost with respect to the control is a simple function of the adjoint state defined by (2.4):

$$\frac{\mathcal{D}\mathcal{J}_2}{\mathcal{D}\Phi} = \ell^2\,\Phi + B_2^*\,\lambda. \tag{2.5}$$

2.1.6. *Solution of control problem*

The most suitable control on $(0, T]$ which results in

$$\frac{\mathcal{D}\mathcal{J}_2}{\mathcal{D}\Phi} = 0 \quad \text{(minimum)} \tag{2.6}$$

may *not* be found simply by setting the gradient $\mathcal{D}\mathcal{J}_2/\mathcal{D}\Phi$ in (2.5) equal to zero, as this gradient information is accurate only in a small neighborhood of the reference solution upon which the matrices A and B_2 were based.† Instead, a more stable iterative approach is used based on the gradient vector:

$$\Phi^{k+1} = \Phi^k - \alpha\frac{\mathcal{D}\mathcal{J}_2}{\mathcal{D}\Phi}, \tag{2.7}$$

where k indicates the iteration index. Thus, the condition (2.6) is approached iteratively according to the following procedure:

† One may propose a Newton-Raphson technique to determine the control, setting the local expression for $\mathcal{D}\mathcal{J}_2/\mathcal{D}\Phi$ in (2.5) equal to zero to determine a new control, determining a new reference condition from (2.1), examining the new perturbation problem to determine a new expression for $\mathcal{D}\mathcal{J}_2/\mathcal{D}\Phi$, and iterating until convergence. However, such a technique is not recommended, as there is no way to insure that the initial reference condition is sufficiently close to a minimum to guarantee convergence of this approach.

(*a*) Initialize control Φ on $(0,T]$ to $\Phi = 0$.
(*b*) Determine state $\mathbf{u}$ on $(0,T]$ from state equation (2.1).
(*c*) Determine adjoint λ on $[0,T)$ from adjoint equation (2.4).
(*d*) Determine local expression for gradient $\mathscr{D}\mathcal{J}_2/\mathscr{D}\Phi$ from (2.5).
(*e*) Test various different values for α in (2.7), computing the resulting state $\mathbf{u}$ from (2.1) and the resulting cost $\mathcal{J}_2$ from (2.2), and determine via a line minimization algorithm that value of α which results in the smallest $\mathcal{J}_2$.
(*f*) Update control Φ on $(0,T]$ via (2.7) using best value of α determined in step *e*.
(*g*) Repeat from step *b* until convergence.

2.2. *Robust regulation of nonlinear problems*

2.2.1. *State equation*

Consider the nonlinear state equation of (2.1) with additional forcing due to an external disturbance χ

$$\boxed{\dot{\mathbf{u}} = \mathcal{A}(\mathbf{u}) + \mathcal{B}_1(\chi) + \mathcal{B}_2(\Phi)} \tag{2.8a}$$

with given initial conditions

$$\mathbf{u} = \mathbf{u}(0) \qquad \text{at } t = 0. \tag{2.8b}$$

The nonlinear function $\mathcal{B}_1(\chi)$ may itself be a function of time.

2.2.2. *Cost function*

The object of applying control in the present case is identical to the robust linear regulation problem described in §1.2.2. Mathematically, it is expressed as the minimization of a cost function $\mathcal{J}_\infty$ with respect to the control Φ while simultaneously maximizing $\mathcal{J}_\infty$ with respect to the disturbance χ, where

$$\boxed{\mathcal{J}_\infty \equiv \frac{1}{2\,T} \int_0^T \left(\mathbf{u}^* \, C_1^* \, C_1 \, \mathbf{u} + \ell^2 \, \Phi^* \, \Phi - \gamma^2 \chi^* \, \chi \right) \, dt}$$

A technique to determine the control Φ on the interval $(0,T]$ which (locally) minimizes the cost function $\mathcal{J}_\infty$ in the presence of a small component of the worst external disturbance χ forcing the state equation (2.8) is now briefly outlined.

2.2.3. *Perturbation equation*

Consider the linear problem of a small perturbation $(\Phi', \chi', \mathbf{u}')$ to some reference solution $(\Phi, \chi, \mathbf{u})$ of the system given by (2.8). It is easily shown that such a perturbation must obey a linear evolution equation of the form

$$\boxed{\dot{\mathbf{u}}' = A\,\mathbf{u}' + B_1\,\chi' + B_2\,\Phi'} \tag{2.9a}$$

with initial conditions

$$\mathbf{u}' = 0 \qquad \text{at } t = 0. \tag{2.9b}$$

The matrices A, B_1, and B_2 are functions of time and depend explicitly on the reference condition $(\Phi, \chi, \mathbf{u})$.

2.2.4. *Adjoint equation*

An adjoint system is defined based on the A matrix of the perturbation problem (2.9) such that

$$-\dot{\lambda} = A^* \lambda + C_1^* C_1 \mathbf{u} \tag{2.10a}$$

with initial conditions

$$\lambda = 0 \qquad \text{at } t = T. \tag{2.10b}$$

2.2.5. *Gradients of cost function $\mathcal{J}_\infty$ with respect to control Φ and disturbance χ*

As in the linear case, the gradients of the cost with respect to the control Φ and the disturbance χ are simple functions of the adjoint state defined by (2.10):

$$\frac{\mathscr{D}\mathcal{J}_\infty}{\mathscr{D}\Phi} = \ell^2 \Phi + B_2^* \lambda \qquad \text{and} \qquad \frac{\mathscr{D}\mathcal{J}_\infty}{\mathscr{D}\chi} = -\gamma^2 \chi + B_1^* \lambda. \tag{2.11}$$

2.2.6. *Solution of control problem*

The most suitable control and disturbance which result in

$$\frac{\mathscr{D}\mathcal{J}_\infty}{\mathscr{D}\Phi} = 0 \ \text{(minimum)} \qquad \text{and} \qquad \frac{\mathscr{D}\mathcal{J}_\infty}{\mathscr{D}\chi} = 0 \ \text{(maximum)} \tag{2.12}$$

may not, strictly speaking, be found simply by setting the gradients $\mathscr{D}\mathcal{J}_\infty/\mathscr{D}\Phi$ and $\mathscr{D}\mathcal{J}_\infty/\mathscr{D}\chi$ in (2.11) equal to zero, as this gradient information is accurate only in a small neighborhood of the reference solution upon which the matrices A, B_1, and B_2 were based. Instead, an iterative approach is used based on the gradient vectors:

$$\Phi^{k+1} = \Phi^k - \alpha \frac{\mathscr{D}\mathcal{J}_\infty}{\mathscr{D}\Phi} \qquad \text{and} \qquad \chi^{k+1} = \chi^k + \beta \frac{\mathscr{D}\mathcal{J}_\infty}{\mathscr{D}\chi} \tag{2.13}$$

where k indicates the iteration index. The iteration procedure followed is analagous to that described in §2.1.6; in the present case, a value of α is chosen to reduce $\mathcal{J}_\infty$ while simulataneously a value of β is chosen to increase $\mathcal{J}_\infty$. The min/max problem is solved when the conditions given in (2.12) are approached.

2.2.7. *Approximate solution for systems of very large dimension*

The min/max problem described by (2.13) is unfeasible when the state equation (2.8) is a model of turbulent channel flow, as the state $\mathbf{u}$ upon which the disturbance acts in this case, and therefore any general representation of the disturbance χ itself, has a very large dimension ($\mathcal{O}(10^7)$ at $Re_\tau = 180$). Thus, instead of forcing the state equation with a disturbance χ determined by the iterative approach given in (2.13), which is guaranteed to be stable but would present excessive computational storage requirements, we settle on a simpler, though possibly unstable, approach for the determination of χ.† To do this, we choose the noise by setting $\mathscr{D}\mathcal{J}_\infty/\mathscr{D}\chi$ in (2.11) equal to zero to determine the the disturbance χ. Taking the matrix B_1 as simply the identity matrix, the disturbance determined in this fashion is proportional to the adjoint field itself

$$\chi = \frac{1}{\gamma^2} \lambda \tag{2.14}$$

† Note that we still determine the control Φ via the stable iterative approach given in (2.13).

For sufficiently large γ (*i.e.*, sufficiently small level of noise), this is an accurate approximation of the global maximum $\mathscr{D}\mathcal{J}_\infty/\mathscr{D}\chi = 0$, and thus results in an accurate approximation of the "worst case" noise. For smaller values of γ (*i.e.*, larger noise levels), this approach can not even be guaranteed to be stable. Trial and error will indicate for what values of γ this approach converges.

REFERENCES

BEWLEY, T.R., AGARWAL, R., & LIU, S. 1996 Optimal and robust control of transition. Submitted to *J. Fluid Mech.*

BEWLEY, T.R., MOIN, P., & TEMAM, R. 1997 Optimal control of turbulence. Under preparation for submission to *J. Fluid Mech.*

DOYLE, J.C., GLOVER, K., KHARGONEKAR, P.P., & FRANCIS, B.A. 1989 State-Space Solutions to Standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Control Problems *IEEE Trans. Auto. Control* **34**, 8, 831-847.

PART D.

Optimization of practical feedback rules for turbulence control

A new method based on control theory for optimizing feedback control rules with the objective of reducing drag in wall-bounded turbulent flows is presented. Both linear and nonlinear control rules (of the type commonly used in neural networks) are considered. These control rules relate wall measurements of skin friction and pressure to the control, which is applied as a continuous distribution of wall-normal boundary velocity with zero net transpiration. Though the optimization technique itself requires complete information about the flow, and thus can only be performed computationally, it is intended that the resulting optimized rules be scaled appropriately and used in physical boundary layer control implementations.

Using optimal control theory, the sensitivity of some representative cost functional to small modifications in the coefficients of a feedback control rule are found via the solution of an adjoint problem. With this sensitivity field, the coefficients are iteratively updated with a gradient algorithm until the cost functional is minimized. Given that this optimization is performed in a representative situation, the coefficients then may be fixed and the control rule effectively used in other flows with similar configurations, requiring only information about the flow which can be obtained with flush-mounted sensors on the wall.

---

1. Introduction

Optimal control theory applied to turbulence provides a rigorous framework to determine the gradient of a cost functional (which represents a physical problem of interest) with respect to small modifications of the control forcing (Bewley, Temam, and Moin, 1997). With such information, combined with a gradient algorithm to update the control, very effective control distributions may be determined. Numerical simulations of this approach in a low Reynolds number turbulent channel flow obtained a 50% drag reduction and an order of magnitude turbulent kinetic energy reduction with small levels of boundary velocity control. Important drawbacks of this approach, however, are 1) it requires complete information about the turbulent fluctuations in the near-wall region, and 2) it is extremely computationally expensive. Thus, it is impossible to apply the optimal control approach directly in an experimental setting.

In order to arrive at a practical scheme, a method was sought to optimize control rules which 1) require only flow information obtainable with wall-mounted sensors, and 2) are computationally inexpensive enough to apply in real time. Possible approaches for this purpose can be divided into two broad categories: *state trajectory approaches*, which attempt to drive some description of the turbulent state (or a portion thereof) in a desired manner, and *direct approaches*, which bypass any description of the turbulent state *per se*, but simply seek a control rule which achieves a desired effect, such as the reduction of drag.

As an example of one state trajectory approach, an adaptive inverse technique has

been applied to a low Reynolds number turbulent channel flow, providing approximately 18% drag reduction (Kim, 1996). This approach first develops an approximate "inverse" model between measurable flow quantities (as input) and the control forcing (as output) with an adaptive technique. Each iteration of the adaptation consists of three steps: 1) computing the error of the model output with respect to the desired model output (the actual control forcing used), 2) determining the influence of the weights in the model on this error, then 3) updating all the weights in the model a small amount in a manner that reduces the error. In neural networks, this is commonly referred to as "back-propagation" of the error. Once this approximate inverse model between the flow measurements and the control converges, the inverse model is used to compute a control which will drive the flow measurements to some desired state. In the case of Kim (1996), the desired state is chosen to be a state with reduced spanwise drag fluctuations.

Drawbacks of the adaptive inverse approach are 1) an *ad hoc* desired state must be chosen, 2) a random "dither" signal needs to be applied to the control in order for the inverse model to have "persistently exciting" data from which to learn, which reduces the performance of the controller, and 3) it is possible that even at statistical steady state, due to the nonlinear nature of the Navier-Stokes equation, the weights in the inverse model may need to continually adapt in order to represent a temporally evolving relationship between the flow measurements and the control. Thus, if the training of the inverse model does not converge fast enough, it will not have time to keep up with the temporal evolution of the flow (for instance, the movement of the near-wall turbulent coherent structures), and may not develop an accurate model between flow measurements and the control which produces them.

Other state trajectory approaches attempt to control a more complete description of the turbulence using a low-dimensional (10-20 mode) representation of the near-wall coherent structures (Coller *et al.*, 1994). In this approach, the orbit of the near wall structures in this representation is partially stabilized, resulting in a reduced "bursting" frequency and, presumably, reduced drag. Coller *et al.* (1994) showed that the frequency of bursting events could be reduced in their model equations, but did not demonstrate how effective such an approach would be at reducing drag when applied to a fully turbulent flow.

Drawbacks of this low-dimensional representation approach include 1) an accurate estimation of this low-dimensional state needs to be made from the measurements at the wall, and 2) a desired *ad hoc* state trajectory must be chosen, which can only be selected well if one has a detailed understanding of the cause/effect relationship of the drag-producing phenomena in the near-wall region, which is still under debate.

Direct approaches may be proposed which bypass estimation and control of the state trajectory altogether. In such approaches, one simply represents the control objective mathematically as a cost functional, then attempts to find a control rule which minimizes this functional.

The simplest direct approach is an adaptive "reinforcement learning" approach. In such an approach, the weights of a control rule are initialized randomly and the control rule applied to the flow. Every time a "good" result is seen (*e.g.*, the drag is reduced), the weights contributing most to the control at that instant are increased, and every time a "bad" result is seen, the corresponding responsible weights are decreased.

The main drawback of this approach, however, is that this reward/punish training algorithm is not very reliable, especially for complicated nonlinear systems, and thus the scheme may not converge at all.

Thus, we arrive at the motivation for the current work, in which we derive a rigorous algorithm to efficiently optimize a direct control scheme, with the goal in mind simply

of reducing some integral measure of the control objective without the prescription of a desired state trajectory. This approach, based on computation of the gradient of a cost functional with respect to modification of the weights in the control rule, will be outlined in the following sections. Numerical simulations that implement this technique are underway.

2. Problem Statement

Our goal is to determine a relationship which takes as input the measurable flow quantities and produces as output a control Φ (the normal component of velocity at the wall) which effectively controls the flow system. The measurable flow quantities are taken to be the wall values of the streamwise drag $\mu\,\partial u_1/\partial n$, the pressure p, and the spanwise drag $\mu\,\partial u_3/\partial n$, where u_i are the components of the velocity, μ is the viscosity, and n_i is a unit normal on each wall facing *into* the flow. Hence, one of the simplest (linear) control rules which may be proposed for this purpose is

$$\Phi = w_1 * \mu\frac{\partial u_1}{\partial n} + w_2 * p + w_3 * \mu\frac{\partial u_3}{\partial n}. \tag{2.1a}$$

The task at hand is simply to optimize the weighting functions w_κ such that the control Φ determined by linear relationship (2.1a) effectively controls the flow system. This configuration is illustrated graphically in Figure 1. Note that the w_κ are convolution functions, where the convolution is defined such that, for example,

$$w_1 * \mu\frac{\partial u_1}{\partial n} = \int_\Gamma w_1(\bar{x})\ \mu\frac{\partial u_1}{\partial n}(x-\bar{x})\ d\bar{\Gamma}, \tag{2.1b}$$

where Γ is the portion of the (2D) boundary of the (3D) flow domain Ω over which measurements are made (and also, we assume, the portion of the boundary over which control is applied), and $\bar{x} \in \bar{\Gamma}$ is the variable of integration. By optimizing the convolution functions $w_\kappa(\bar{x})$, we take into account "nearby" flow measurements (in the direction $\bar{x}$) from a specific actuator location (x). In fact, the extent to which these convolution functions are nonzero when converged will indicate how far in each direction from a specific actuator flow measurements are relevant when computing an effective control. Additionally, we will constrain the spatial average of each convolution function to be zero, so that the net control on each wall is exactly zero at any instant. This is motivated both by physical flow control devices and the current simulations which require control with zero net mass flux. We will seek the best control rule to interact with the fluctuating part of the turbulence only.

Note also that the weighting functions w_κ are prescribed at the outset to be *invariant in time*. Though the method used requires that the weights be optimized by considering finite time horizon $[0,T]$, we seek to approximate the steady-state weights at the "infinite time horizon" in which turbulent fluctuations near the wall are countered by a fixed control rule at the wall in an efficient drag-reducing manner.

As a straightforward extension to this work, one may also optimize nonlinear control rules of a form similar to that used in neural networks, which may be written

$$\Phi = \sum_{\lambda=1}^{\Lambda} W_\lambda\, g(\xi_\lambda),$$

$$\xi_\lambda = w_{1\lambda} * \mu\frac{\partial u_1}{\partial n} + w_{2\lambda} * p + w_{3\lambda} * \mu\frac{\partial u_3}{\partial n} + B_\lambda,$$

where Λ is the number of "nodes" in the "hidden layer" of the network and $g(\xi_\lambda)$ is an

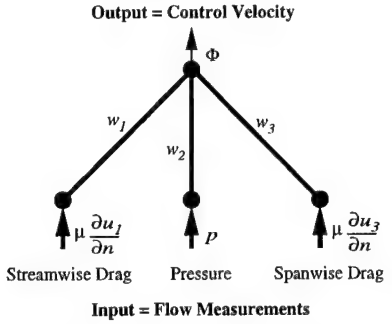


FIGURE 1. Linear Network. The flow measurements which we take as inputs are localized measurements on the wall of the streamwise drag, the pressure, and the spanwise drag. The flow measurements are convolved with the weighting functions w_κ and summed to determine the control Φ. The input flow measurements are field variables and are indicated with heavy lines—the corresponding weights are convolution functions (in the continuous case) or two dimensional arrays containing a stencil of weights (in the discrete case).

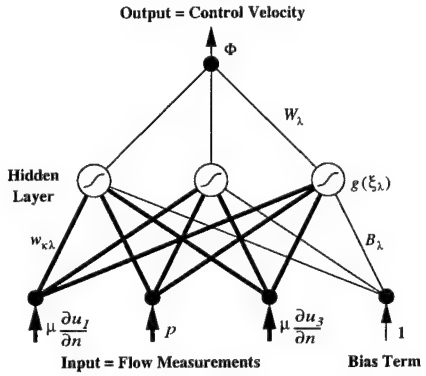


FIGURE 2. Nonlinear Network. The output of several simple networks λ similar to the one depicted in Figure 1 (with added bias weights B_λ connected to an input clamped to unity) are used as the arguments ξ_λ to activation functions g at the hidden nodes. The output of all of the hidden nodes $g(\xi_\lambda)$ are then weighted with the W_λ and summed to produce the control Φ.

"activation function" which will be prescribed. Control rules of this form, used commonly in neural networks, have seen a broad range of application and are capable of representing very general nonlinear relationships (Hertz *et al.*, 1991). The task at hand in this case is to optimize the weighting functions $w_{\kappa\lambda}$ and the discrete weights B_λ and W_λ such that Φ effectively controls the flow system. This type of network is illustrated in Figure 2.

For simplicity, the equations necessary to optimize the linear network of equation (2.1) and figure 1 will be derived in this paper.

3. Governing equations

The flow system we consider is fully developed turbulent channel flow with periodic boundary conditions in the streamwise (x_1) and spanwise (x_3) directions, as shown in Figure 3. Blowing and suction through the computational equivalent of holes drilled in the wall will be applied according to the linear control rule (2.1). The control rule

FIGURE 3. Flow configuration. Blowing and suction is applied through closely spaced holes drilled in the walls to control the flow. This control will be coordinated with nearby flow measurements on the wall in a manner to be determined.

optimized for this configuration should also be effective in turbulent boundary layers due to the similar near-wall behavior of these flows.

The governing equations may be written functionally as

$$\mathcal{N}(U) = 0 \qquad \text{in } \Omega \tag{3.1a}$$

with boundary conditions

$$u_i = \Phi n_i \qquad \text{on } \partial\Omega_1, \tag{3.1b}$$

where Φ is determined by the linear control rule

$$\Phi = w_1 * \mu \frac{\partial u_1}{\partial n} + w_2 * p + w_3 * \mu \frac{\partial u_3}{\partial n}, \tag{3.1c}$$

and prescribed initial conditions

$$u_i = u_i(0) \qquad \text{at } t = 0. \tag{3.1d}$$

For clarity, all differential equations will be written in operator form in this discussion, with these operators defined when first introduced. The (nonlinear) Navier-Stokes operator for the present case, in which the flow is assumed to have uniform density and viscosity, is given by

$$\mathcal{N}(U) = \begin{pmatrix} \dfrac{\partial u_i}{\partial t} + u_j \dfrac{\partial u_i}{\partial x_j} - \nu \dfrac{\partial^2 u_i}{\partial x_j^2} + \dfrac{1}{\rho}\dfrac{\partial p}{\partial x_i} + \dfrac{1}{\rho}\delta_{1i} P_x \\ \dfrac{\partial u_j}{\partial x_j} \end{pmatrix}.$$

In this discussion, x_1 is the streamwise direction, x_2 is the wall-normal direction, x_3 is the spanwise direction, the u_i's are the corresponding velocities, p is the pressure, ρ is the density, μ is the absolute viscocity, and $\nu \equiv \mu/\rho$ is the kinematic viscocity. The flow is sustained by a mean pressure gradient P_x in the streamise direction, which is modified at each time step in order to maintain a constant mass flux through the channel. Define also n as a wall-normal unit vector directed *into* the channel, $\overline{\tau_w} \equiv \mu\, \partial\overline{u_1}/\partial n|_{wall}$ as the mean skin friction on the wall for the uncontrolled channel (averaged in space and time), $u_\tau \equiv (\overline{\tau_w}/\rho)^{1/2}$ as the mean friction velocity, δ as the channel half-width, and $Re_\tau \equiv u_\tau \delta/\nu$ as the Reynolds number based on the mean friction velocity and the channel half width. The flow considered in this work is taken at $Re_\tau = 180$. All velocities are normalized by the friction velocity u_τ, and therefore may also be marked with a $(^+)$ superscript. All lengths are normalized by δ unless marked with a $(^+)$ superscript, in

which case they are normalized by the wall unit ν/u_τ. All times are normalized by δ/u_τ unless marked with a $(^+)$ superscript, in which case they are normalized by ν/u_τ^2.

We will now develop a systematic procedure to optimize the weighting functions w_κ. The first step is to define the control problem of interest mathematically as a cost functional to be minimized.

4. Cost functional

The objective of applying the control in this problem is to reduce the drag without using excessive amounts of control forcing. Mathematically, a cost functional for this problem may thus be expressed as

$$\mathcal{J}(w) = \frac{1}{AT} \int_\Gamma \int_0^T \left(\mu \frac{\partial u_1}{\partial n} + \frac{\ell^2}{2} \Phi^2 \right) dt \, d\Gamma. \tag{4.1}$$

The term involving $\mu \, \partial u_1/\partial n$ is the drag averaged over the wall Γ and the time interval of interest $[0, T]$. The term involving Φ^2 is an expression of the magnitude of the control. These two terms are weighted together with a factor ℓ^2, which represents the price of the control. This quantity is small if the control is "cheap", and large if applying the control is "expensive".

Minimization of $\mathcal{J}$ corresponds to reducing drag while maintaining a small amount of control forcing.

5. Gradient of cost functional

We now develop a technique to compute the gradient of the cost functional $\mathcal{J}$ with respect to the weighting functions w.

5.1. *Perturbation field*

Consider first the Fréchet differential of the flow U with respect to w, which is defined such that

$$\begin{aligned} \dot{U} &\equiv \frac{1}{A} \lim_{\epsilon \to 0} \frac{U(w + \epsilon \dot{w}) - U(w)}{\epsilon} \\ &= \frac{1}{A} \int_\Gamma \sum_{\kappa=1}^{3} \frac{\mathscr{D} U(w)}{\mathscr{D} w_\kappa} \dot{w}_\kappa \, d\Gamma \end{aligned}$$

where $\dot{w}$ is an arbitrary "update direction" to the weighting function w. This update direction will remain undetermined and will later be isolated and removed from the equation for the perturbation of the cost functional caused by a perturbation to the control. The perturbation field $\dot{U}$ is governed by the Fréchet differential of (3.1) with respect to w, which may be written:

$$\mathcal{A} \dot{U} = 0 \qquad \text{in } \Omega \tag{5.1a}$$

with boundary conditions

$$u_i = \dot{\Phi} \, n_i \qquad \text{on } \partial\Omega_1, \tag{5.1b}$$

where

$$\dot{\Phi} = w_1 * \mu \frac{\partial \dot{u}_1}{\partial n} + w_2 * \dot{p} + w_3 * \mu \frac{\partial \dot{u}_3}{\partial n} + \dot{w}_1 * \mu \frac{\partial u_1}{\partial n} + \dot{w}_2 * p + \dot{w}_3 * \mu \frac{\partial u_3}{\partial n}, \tag{5.1c}$$

and with initial conditions

$$\dot{u}_i = 0 \qquad \text{at } t = 0. \tag{5.1d}$$

The Fréchet differential of the (non-linear) Navier-Stokes operator is given by

$$\mathcal{A}\dot{U} = \begin{pmatrix} \dfrac{\partial \dot{u}_i}{\partial t} + u_j \dfrac{\partial \dot{u}_i}{\partial x_j} + \dot{u}_j \dfrac{\partial u_i}{\partial x_j} - \nu \dfrac{\partial^2 \dot{u}_i}{\partial x_j^2} + \dfrac{1}{\rho} \dfrac{\partial \dot{p}}{\partial x_i} \\ -\dfrac{1}{\rho} \dfrac{\partial \dot{u}_j}{\partial x_j} \end{pmatrix}, \tag{5.2}$$

which is linear in the perturbation field $\dot{U}$, but is a function of the solution U of the Navier-Stokes problem, so that $\mathcal{A} = \mathcal{A}(U)$.

The Fréchet differential of the cost functional $\mathcal{J}$ with respect to w is:

$$\begin{aligned} \dot{\mathcal{J}} &\equiv \frac{1}{A} \lim_{\epsilon \to 0} \frac{\mathcal{J}(w + \epsilon \dot{w}) - \mathcal{J}(w)}{\epsilon} \\ &= \frac{1}{A} \int_\Gamma \sum_{\kappa=1}^{3} \frac{\mathscr{D}\mathcal{J}(w)}{\mathscr{D} w_\kappa} \dot{w}_\kappa \, d\Gamma \\ &= \frac{1}{AT} \int_\Gamma \int_0^T \left(\mu \frac{\partial \dot{u}_1}{\partial n} + \ell^2 \, \Phi \, \dot{\Phi} \right) dt \, d\Gamma \end{aligned}$$

It is seen that the perturbation of the cost functional $\dot{\mathcal{J}}$ is a function of the perturbation of the flow $\dot{U}$. The linear dependence of $\dot{U}$ on $\dot{w}$ may, in theory, be found directly from (5.1). However, in practice, this is not a tractable approach due to the excessively large dimension of the problem under consideration. Thus, we seek a simpler way to express the above equation in a manner in which the gradient $\mathscr{D}\mathcal{J}(w)/\mathscr{D} w_\kappa$ may be determined. It is for this reason that we now propose the definition of an adjoint field.

5.2. *Definition of adjoint field*

As discussed in previous sections, an adjoint operator $\mathcal{A}^*$ may be defined by the identity

$$< \mathcal{A}\dot{U}, \tilde{U} > = < \dot{U}, \mathcal{A}^* \tilde{U} > + b. \tag{5.3}$$

Integration by parts may be used to move all differential operations from $\dot{U}$ on the left hand side of the equation to $\tilde{U}$ on the right hand side, resulting in an expression for the adjoint operator

$$\mathcal{A}^* U^* = \begin{pmatrix} -\dfrac{\partial u_i^*}{\partial t} - u_j \left(\dfrac{\partial u_i^*}{\partial x_j} + \dfrac{\partial u_j^*}{\partial x_i} \right) - \nu \dfrac{\partial^2 u_i^*}{\partial x_j^2} + \dfrac{1}{\rho} \dfrac{\partial p^*}{\partial x_i} \\ -\dfrac{1}{\rho} \dfrac{\partial u_j^*}{\partial x_j} \end{pmatrix}, \tag{5.4}$$

where the operator $\mathcal{A}^*$ is a function of U, and an expression for b, which contains all the boundary terms:

$$b = \int_w \int_0^T -n_j \left(u_j\, u_i'\, u_i^* + u_i\, u_j'\, u_i^* - \nu\Big(\frac{\partial u_i'}{\partial x_j}\, u_i^* - u_i' \frac{\partial u_i^*}{\partial x_j}\Big) + \frac{1}{\rho}\Big(p'\, u_j^* - u_j'\, p^*\Big)\right) dt\, dS$$
$$+ \int_\Omega u_i'\, u_i^*\, dV \Big|_{t=T} - \int_\Omega u_i'\, u_i^*\, dV \Big|_{t=0}.$$

In order to express the perturbation of the cost functional in a usable form, we now define an adjoint state by the system of equations

$$\mathcal{A}^* \tilde{U} = 0, \tag{5.5a}$$

with mixed boundary conditions on the walls

$$\begin{aligned} \tilde{u}_1 \; + \tilde{f} * \check{w}_1 &= 1 + \ell^2\, \Phi * \check{w}_1 \\ -\tilde{u}_2\, n_2 + \tilde{f} * \check{w}_2 &= \quad \ell^2\, \Phi * \check{w}_2 \\ \tilde{u}_3 \; + \tilde{f} * \check{w}_3 &= \quad \ell^2\, \Phi * \check{w}_3, \end{aligned} \tag{5.5b}$$

where

$$\tilde{f} \equiv \tilde{p} - 2\,\rho\, \tilde{u}_2\, u_2 - \mu \frac{\partial \tilde{u}_2}{\partial x_2}$$
$$\check{b}(x) \equiv b(-x),$$

and with initial conditions

$$\tilde{u}_i = 0 \text{ at } t = T. \tag{5.5c}$$

5.3. *Identification of gradient*

Using the identity (5.3) and the definition of the adjoint in (5.5), we can algebraically manipulate (5.3) to the form

$$\frac{1}{A} \int_\Gamma \sum_{\kappa=1}^{3} \frac{\mathscr{D}\mathcal{J}(w)}{\mathscr{D} w_\kappa}\, \dot{w}_\kappa\, d\Gamma = \frac{1}{A} \int_\Gamma \sum_{\kappa=1}^{3} \tilde{G}_\kappa\, \dot{w}_\kappa\, d\Gamma, \tag{5.6}$$

where $\tilde{G}_\kappa$ is some function of the solution to the adjoint problem (5.5). As $\dot{w}$ is arbitrary, we may then identify the expression for $\tilde{G}_\kappa$ as $\mathscr{D}\mathcal{J}(w)/\mathscr{D} w_\kappa$. It may be shown that the resulting expression for the gradient is

$$\frac{\mathscr{D}\mathcal{J}(w)}{\mathscr{D} w_1} = \frac{1}{T} \int_0^T \Big(-\tilde{f} + \ell^2\, \Phi \Big) * \mu \frac{\partial \check{u}_1}{\partial n}\, dt$$
$$\frac{\mathscr{D}\mathcal{J}(w)}{\mathscr{D} w_2} = \frac{1}{T} \int_0^T \Big(-\tilde{f} + \ell^2\, \Phi \Big) * \check{p}\, dt$$
$$\frac{\mathscr{D}\mathcal{J}(w)}{\mathscr{D} w_3} = \frac{1}{T} \int_0^T \Big(-\tilde{f} + \ell^2\, \Phi \Big) * \mu \frac{\partial \check{u}_3}{\partial n}\, dt$$

Thus, the gradient of the cost $\mathcal{J}(w)$ with respect to modification of the weights w of the feedback control rule has been represented as a function of the solution of an adjoint problem. This adjoint problem, though linear, has approximately the same complexity as the Navier-Stokes problem itself, as can be seen by examining Equation (5.5) and the definition of $\mathcal{A}^*$. Note that the adjoint field evolves backwards over the domain $[0, T]$ from initial conditions at $t = T$; this is the natural direction in time to march these

equations due to the sign of the viscous and time derivatives in $\mathcal{A}^*$. Note also that the computation of the adjoint field requires the storage of the flow field over the entire interval $[0,T]$, as the adjoint operator depends on the flow velocity, *i.e.* $\mathcal{A}^* = \mathcal{A}^*(U)$.

6. Gradient update to control

With the gradients computed using the adjoint field, a control rule may be optimized using a gradient algorithm as done in previous sections.

7. Computational Results

At the time this document went to press, computational results were not yet available. They will appear in the version of to be submitted to *J. Fluid Mech.*

8. Discussion

A new technique for optimizing linear and nonlinear feedback control rules for turbulent flows has been presented. This technique is based solely on the equations governing the flow and a mathematical statement of the control objective, thus bypassing the *ad hoc* identification of a desired state trajectory often used to determine feedback control rules. Also, the training is based on an adjoint ("sensitivity") field, which determines the gradient of the cost with respect to small modifications of the weights in a rigorous manner. Thus, convergence can be expected to be much better than for an reinforcement learning approach with an adaptive algorithm.

A straightforward extension of the present approach is to take into account *past* measurements in the control rule. Past measurements, which may easily be stored in an experimental implementation, may give additional information about the convection velocity of flow structures which cannot be determined from instantaneous measurements alone. It is also possible that such information can be determined by recurrent networks, in which the inputs of the control network include the outputs of the network from the previous time step.

Drawbacks of the present method include 1) an accurate mathematical model of the flow equations and boundary conditions are needed for the training, and 2) the training algorithm is quite complex, requiring simulation on a supercomputer. However, this method should provide insight into effective new control rules which one could not think of otherwise, and which can be further modified to fit practical problems. In addition, they may be used to guide the development of experimental configurations, revealing the necessary locations of sensors with respect to the actuators in order to obtain information relevant to effective control strategies.

REFERENCES

ABERGEL, F. & TEMAM, R. 1990 On some control problems in fluid mechanics *Theor. and Comp. Fluid Dynamics* **1**, 303.

BEWLEY, T. 1996 Optimal control of turbulence PhD Thesis (in preparation), Department of Mechanical Engineering, Stanford University.

COLLER, B., HOLMES, P. & LUMLEY, J. 1994 Interaction of adjacent bursts in the wall region *Phys. Fluids* **6**, 954-961.

HERTZ, J., KROGH, A., & PALMER, R. 1990 *Introduction to the theory of neural computation* Addison Wesley.

MOIN, P. & BEWLEY, T. 1995 Application of control theory to turbulence *Twelfth Australasian Fluid Mechanics Conference*, December 10-15, Sydney, Australia, p. 109.